

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 15:07:17 ; Search time 23 Seconds
(without alignments)
445.420 Million cell updates/sec

Title: US-09-762-277A-1

Perfect score: 1327

Sequence: 1 MNPLILIAFLGAATPTDD.....VYTKVCFVDWIQSTIAANS 247

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1327	100.0	247	1	TRY2_CANFA
2	1153	86.9	247	1	TRY2_BOVIN
3	1127	84.9	246	1	TRY2_MOUSE
4	1111	83.7	246	1	TRY2_RAT
5	1101	83.0	246	1	TRY1_RAT
6	1098.5	82.8	247	1	TRY3_RAT
7	1097	82.7	247	1	TRY2_HUMAN
8	1085	81.8	247	1	TRY1_HUMAN
9	1071	80.7	231	1	TRYP_PIG
10	1058	79.7	247	1	TRY3_HUMAN
11	1036	78.1	304	1	TRY4_HUMAN
12	1033	77.8	244	1	TRY2_XENLA
13	1026	77.3	246	1	TRY1_CANFA
14	1019	76.8	243	1	TRY1_BOVIN
15	992.5	74.8	243	1	TRY1_XENLA
16	966.5	72.8	248	1	TRY3_CHICK
17	960.5	72.4	247	1	TRY4_RAT
18	948.5	71.5	238	1	TRY3_SALSA
19	940	70.8	248	1	TRY2_CHICK
20	937	70.6	248	1	TRY1_CHICK
21	905	68.2	246	1	TRYB_RAT
22	896	67.5	246	1	TRYA_RAT
23	877	66.1	242	1	TRY1_SALSA
24	867	65.3	231	1	TRY2_SALSA
25	855.5	64.5	229	1	TRYP_SQUAC
26	834.5	62.9	241	1	TRYX_GADMO
27	824.5	62.1	241	1	TRY1_GADMO
28	649.5	48.9	250	1	TRYP_PLEPL
29	614	46.3	251	1	KLKE_HUMAN
30	578	43.6	293	1	KLK5_HUMAN
31	576	43.4	250	1	KLKB_HUMAN
32	572	43.1	256	1	KLKF_HUMAN
33	570.5	43.0	260	1	KLK8_HUMAN

34	567.5	42.8	244	1	KLK6_HUMAN
35	555	41.8	260	1	NRPN_MOUSE
36	555	41.8	260	1	NRPN_RAT
37	546	41.1	248	1	KLKC_HUMAN
38	538.5	40.6	277	1	KLKD_HUMAN
39	535	40.3	262	1	KLK1_HUMAN
40	534	40.2	250	1	KLK9_HUMAN
41	528.5	39.8	257	1	KLK1_MACFA
42	527.5	39.8	253	1	KLK7_HUMAN
43	526	39.6	258	1	KLK1_PAPHA
44	517.5	39.0	261	1	KLK1_MOUSE
45	508	38.3	261	1	KLK2_HUMAN

ALIGNMENTS

RESULT 1

ID	TRY2_CANFA	STANDARD	PRT	247 AA
AC	P06872;			
DT	01-JAN-1988 (Rel. 06, Created)			
DT	01-JAN-1988 (Rel. 06, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Trypsin, anionic precursor (EC 3.4.21.4).			
OS	Canis familiaris (Dog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.			
OX	NCBI_TaxID=9615;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=86284628; PubMed=3841794;			
RA	Pinsky S.D., Laforge K.S., Scheele G.;			
RT	"Differential regulation of trypsinogen mRNA translation: full-length			
RT	mRNA sequences encoding two oppositely charged trypsinogen isoenzymes			
RT	in the dog pancreas."			
RL	Mol. Cell. Biol. 5:2669-2676(1985).			
CC	- - CATALYTIC ACTIVITY: Preferential cleavage: Arg- -Xaa, Lys- -Xaa.			
CC	- - SUBCELLULAR LOCATION: Extracellular.			
CC	- - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			
CC				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC				
CC	EMBL; M11589; AAA30899.1;			
DR	PIR; A26273; TRDG.			
DR	HSP; P00763; LDPO.			
DR	MEROPS; S01.258; -			
DR	InterPro; IPR001314; Chymotrypsin.			
DR	InterPro; IPR001254; Ser_protease_Try.			
DR	Pfam; PF00089; trypsin; 1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	SMART; SM00020; Tryp_SPC; 1.			
DR	PROSITE; PS00240; TRYP_SIN_DOM; 1.			
DR	PROSITE; PS00134; TRYP_SIN_HIS; 1.			
DR	PROSITE; PS00135; TRYP_SIN_SER; 1.			
KW	Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;			
KW	Multi-gene family.			
FT	SIGNAL 1 15			
FT	PROPEP 16 23			
FT	CHAIN 24 247			
FT	ACT_SITE 63 63			
FT	ACT_SITE 107 107			
FT	ACT_SITE 200 200			
FT	DISULFID 30 160			
FT	DISULFID 48 64			
FT	DISULFID 132 233			
FT	DISULFID 139 206			

Q92876	homo sapien
O61955	mus musculus
O88780	rattus norv
O9ukt0	homo sapien
O9ukt3	homo sapien
P06870	homo sapien
Q9ukt9	homo sapien
Q07276	macaca fasc
P49862	homo sapien
Q28773	papio hamad
P15947	mus musculus
P20151	homo sapien

```

FT DISULFID      171   185          BY SIMILARITY.
FT DISULFID      196   220          BY SIMILARITY.
FT SITE          194   194    REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SQ SEQUENCE     247 AA; 26423 MW; 374E9D31D6DB8EAF CRC64;

Query Match       100.0%; Score 1327; DB 1; Length 247;
Best Local Similarity 100.0%; Pred. No. 1.2e-109;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MNPLLIALFAAAVATPTDDDKKLVGGTYTCSENSVPYQVSLNAGYHFCGGLISDQWVS 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MNPLLIALFAAAVATPTDDDKKLVGGTYTCSENSVPYQVSLNAGYHFCGGLISDQWVS 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      61 AAHCKYSRIQVRLEGEYNIDVLEGNEQFINSAKVIRHPNYNSWIILDNDIMLIKSSPAVLN 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      61 AAHCKYSRIQVRLEGEYNIDVLEGNEQFINSAKVIRHPNYNSWIILDNDIMLIKSSPAVLN 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      121 ARVATISLPRACAPGTQCCLISGWNTLSGSTNYPELLQLCDADPILTOAQCEASYPGOIT 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
--      121 ARVATISLPRACAPGTQCCLISGWNTLSGSTNYPELLQLCDADPILTOAQCEASYPGOIT 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

-/      181 ENMICAGFLGGKDSCOGDSGGPVVCNGELQGIVSWGYGCAQNKNPGVYTKVCNFVDWIQ 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      181 ENMICAGFLGGKDSCOGDSGGPVVCNGELQGIVSWGYGCAQNKNPGVYTKVCNFVDWIQ 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      241 STTAANS 247
| | | | |
Db      241 STTAANS 247
| | | | |

RESULT 2
TRY2_BOVIN ID TRY2_BOVIN STANDARD; PRT; 247 AA.
AC Q29463;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 13-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin, anionic precursor (EC 3.4.21.4).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
ON NCBI_TaxId=9913;
RX [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Holstein-Friesian; TISSUE=Pancreas;
RX MEDLINE=91065383; PubMed=1701147;
-- le Huouon I., Wicker C., Guilletoau P., Toullac R., Puigserver A.;
-- "Isolation and nucleotide sequence of cDNA clone for bovine
-- pancreatic anionic trypsinogen. Structural identity within the
-- trypsin family."
Rt Eur. J. Biochem. 193:767-773(1990).
RL CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-Xaa, Lys-Xaa.
CC CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. Usage by and for commercial
CC or send an email to license@isb-sib.ch.
-----
DR EMBL; X54703; CAA38513.1; -.
DR HSP; P00763; IDPO.
DR MEROPS; S01.258; -.
DR InterPro; IPRO01314; Chymotrypsin.
DR DR InterPro; IPRO01254; Ser.protease_Try.
DR Pfam; PF00089; trypsin_1
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.

```

entitles requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

```
CC EMBL; X04574; CAA28243.1; -
CC EMBL; X04577; CAA28245.1; -
DR PIR; B25528; B25528.
DR HSP; P00763; IDPO.
DR MEROPS; S01.064; -
DR MGI; 102759; TRY2.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT SIGNAL 1 15
FT PROPEP 16 23
FT CHAIN 24 246
FT ACT_SITE 63 63
FT ACT_SITE 107 107
FT ACT_SITE 200 200
FT DISULFID 30 160
FT DISULFID 48 64
FT DISULFID 132 233
FT DISULFID 139 206
FT DISULFID 171 185
FT DISULFID 196 220
FT SITE 194 194
SQ SEQUENCE 246 AA; 26203 MW; CEF8C97AAC2D07AD CRC64;

Query Match 84.98; Score 1127; DB 1; Length 246;
Best Local Similarity 82.58; Pred. No. 4.66-92;
Matches 203; Conservative 20; Mismatches 23; Indels 0; Gaps 0;

QY 1 MNPILLALFAAATPTDDDKIVGGYTCENSVYQVSLNAGYHFCGSLSDQWVNS 60
Db 1 MSALLILALVGAAPVAVFDDDDKIVGGYTCRESSVPYQVSLNAGYHFCGSLINDQWVNS 60

QY 61 AAHCYKRIQVRLGEYINIDYLEGNEQFINSAKVIHRPNYNSWILDNDIMLIKLSPPAVLN 120
Db 61 AAHCYKRIQVRLGEYINIDYLEGNEQFINSAKVIHRPNYNSWILDNDIMLIKLSPPAVLN 120

QY 121 ARVATISLPACAPAGTGQCLISGNGNTLSSGTNPPELLQCLDAPILFOACEASYPGOIT 180
Db 121 ARVATISLPACAPAGTGQCLISGNGNTLSSGTNPPELLQCLDAPILFOACEASYPGOIT 180

QY 121 ARVASPLPSCAPAGTGQCLISGNGNTLSSGTNPPELLQCLDAPILFOACEASYPGDIT 180
Db 121 ARVASPLPSCAPAGTGQCLISGNGNTLSSGTNPPELLQCLDAPILFOACEASYPGDIT 180

QY 181 ENMICAGFLGGKDCQSGDGGPVVNCNGELQGVSWGCGAOKNKPQVYTKVCFNVDWVQ 240
Db 181 ENMICAGFLGGKDCQSGDGGPVVNCNGELQGVSWGCGAOKNKPQVYTKVCFNVDWVQ 240

QY 241 STIAAN 246
Db 241 NTIADN 246

RESULT 4
TRY2_RAT
ID TRY2_RAT STANDARD; PRT; 246 AA.
AC P00763;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin II, anionic precursor (EC 3.4.21.4) (Pretrypsinogen II).
GN TRY2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RX RA MEDLINE=85054880; PubMed=6094547;
RA Craik C.S., Choo Q.L., Swift G.H., Quinto C., McDonald R.J.,
RA Rutter W.J.;
RT "Structure of two related rat pancreatic trypsin genes.";
RL J. Biol. Chem. 259:14255-14264(1984).
RN [2]
RP SEQUENCE OF 9-246 FROM N.A.
RX STRAIN=Sprague-Dawley; TISSUE=Pancreas;
RX MEDLINE=82265624; PubMed=6896710;
RA McDonald R.J., Stary S.J., Swift G.H.;
RT "Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide
RT sequences of the cloned cDNAs.";
RL J. Biol. Chem. 257:9724-9732(1982).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.59 ANGSTROMS).
RX MEDLINE=91351998; PubMed=1881877;
RA Earnest T., Fauman E., Craik C.S., Stroud R.;
RT "1.59-A structure of trypsin at 120 K: comparison of low temperature
RT and room temperature structures.";
RL Proteins 10:171-187(1991).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=96214506; PubMed=8634241;
RA Brinen L.S., Willett W.S., Craik C.S., Fletterick R.J.;
RT "X-ray structures of a designed binding site in trypsin show metal-
RT dependent geometry.";
RL Biochemistry 35:5999-6009(1996).
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
```

```
EMBL; V01274; CAA24581.1; -
EMBL; L00131; AAA98517.1; -
EMBL; L00130; AAA98517.1; JOINED.
PIR; A22657; TRRT2.
DR PDB; 1ANB; 01-APR-97.
DR PDB; 1ANC; 01-APR-97.
DR PDB; 1ANE; 01-APR-97.
DR PDB; 1AMH; 24-DEC-97.
DR PDB; 1SLU; 11-JUL-96.
DR PDB; 1SLV; 11-JUL-96.
DR PDB; 1SLW; 11-JUL-96.
DR PDB; 1SLX; 11-JUL-96.
DR PDB; 1DPO; 07-JUL-97.
DR PDB; 3TGI; 23-DEC-98.
DR PDB; 3TGI; 23-DEC-98.
DR MEROPS; S01.258; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW Multigene family; 3D-structure.
FT SIGNAL 1 15
FT PROPEP 16 23
FT CHAIN 24 246
FT ACT_SITE 63 63
FT ACT_SITE 107 107
FT ACT_SITE 200 200
FT DISULFID 30 160
```



```

FT STRAND 194 194
FT TURN 197 198
FT TURN 200 201
FT STRAND 203 206
FT TURN 207 208
FT STRAND 209 216
FT TURN 223 224
FT STRAND 227 231
FT HELIX 232 234
FT HELIX 236 244
FT TURN 245 245
SQ SEQUENCE 246 AA; 25959 MW; 6AFA0ADAD11943FB5 CRC64;

Query Match 83.0%; Score 1101; DB 1; Length 246;
Best Local Similarity 80.5%; Pred. No. 8.9e-90;
Matches 198; Conservative 26; Mismatches 22; Indels 0; Gaps 0;

QY 1 MNPELLILAFGAATVPTDDDDKIVGGYTCSENSVPYQVSLNAGYHFCGSLSDQWVVS 60
Db 1 MSALLILALVGAATVPTDDDDKIVGGYTCPEHSVPYQVSLNAGYHFCGSLSDQWVVS 60
J 61 AAHCYKSRIQVRLGEYINIDVLENGEQFINSKVRHNPYNYSWILDNDIMLIKLSPPAVLN 120
Db 61 AAHCYKSRIQVRLGEYINIDVLENGEQFINSKVRHNPYNYSWILDNDIMLIKLSPPAVLN 120
QY 121 ARVATISLPACAAAGTQCLISGNGNTLSSTNTYPELLOCLDAPILTOAOCEASYPGQIT 180
Db 121 ARVATISLPACAAAGTQCLISGNGNTLSSTNTYPELLOCLDAPILTOAOCEASYPGQIT 180
QY 181 ENMICAGFLEGGKDCOGDGGPVCNGELQGVSWGVCAGKAKKPGYTKVCFVDFWQI 240
Db 181 SSMICVGFLEGGKDCOGDGGPVCNGELQGVSWGVCAGKAKKPGYTKVCFVDFWQI 240
QY 241 STIAAN 246
Db 241 DTIAAN 246

RESULT 6
TRY3_RAT
ID TRY3_RAT STANDARD; PRT; 247 AA.
AC P08426;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin III, cationic precursor (EC 3.4.21.4) (Pretrypsinogen III).
GN TRY3.
OS Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
MEDLINE=87271609; PubMed=3607011;
RA Fletcher T.S., Alhadeff M., Cralk C.S., Largman C.;
RT "Isolation and characterization of a cDNA encoding rat cationic
trypsinogen."
RL Biochemistry 26:3081-3086(1987).
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -|- SUBCELLULAR LOCATION: Extracellular.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M16624; AAA41985.1;
CC PIR; A27547; A27547.
CC HSSP; P00763; LDPO.

```

```

DR MEROPS; S01.056;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW Multigene family.
FT SIGNAL 1 15
FT PROPEP 16 24
FT CHAIN 25 247
FT ACT_SITE 64 64
FT ACT_SITE 108 108
FT ACT_SITE 201 201
FT DISULFID 31 161
FT DISULFID 49 65
FT DISULFID 133 234
FT DISULFID 140 207
FT DISULFID 172 186
FT DISULFID 197 221
FT SITE 195 195
SQ SEQUENCE 247 AA; 26269 MW; D74892BAA584BA8 CRC64;

Query Match 82.8%; Score 1098.5; DB 1; Length 247;
Best Local Similarity 78.5%; Pred. No. 1.5e-89;
Matches 194; Conservative 31; Mismatches 21; Indels 1; Gaps 1;

QY 1 MNPELLILAFGAATVPTDDDDKIVGGYTCSENSVPYQVSLNAGYHFCGSLSDQWVVS 59
Db 1 MKALIFLAFGAATVPTDDDDKIVGGYTCOKNSLPYQVSLNAGYHFCGSLSDQWVVS 60
QY 60 SAAHCYKSRIQVRLGEYINIDVLENGEQFINSKVRHNPYNYSWILDNDIMLIKLSPPAVLN 119
Db 61 SAAHCYKSRIQVRLGEYINIDVLENGEQFINSKVRHNPYNYSWILDNDIMLIKLSPPAVLN 120
QY 120 NARVATISLPACAAAGTQCLISGNGNTLSSTNTYPELLOCLDAPILTOAOCEASYPGQI 179
Db 121 NSRVSTVSLPRSCGSGGPKCLVSGWNTLSSTNTYPELLOCLDAPILTOAOCEASYPGKI 180
QY 180 TENMICAGFLEGGKDCOGDGGPVCNGELQGVSWGVCAGKAKKPGYTKVCFVDFWQI 239
Db 181 TSNMFCGLFLEGGKDCOGDGGPVCNGELQGVSWGVCAGKAKKPGYTKVCFVDFWQI 240
QY 240 QSTIAAN 246
Db 241 QQTVAAN 247

RESULT 7
TRY2_HUMAN
ID TRY2_HUMAN STANDARD; PRT; 247 AA.
AC P07478;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin II precursor (EC 3.4.21.4) (Anionic trypsinogen).
GN PRS2 OR TRY2 OR TRYP2.
OS Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
MEDLINE=86221712; PubMed=3011602;
RA Emi M., Nakamura Y., Ogawa M., Yamamoto T., Nishide T., Mori T.,
RA Matsubara K.;
RT "Cloning, characterization and nucleotide sequences of two cDNAs
encoding human pancreatic trypsinogens."
RL Gene 41:305-310(1986).
RN [2]

```

ID TRYL_HUMAN STANDARD; PR1: .247 AA.
 AC P07477; Q92955; Q9HAN4; Q9HAN6; Q9HAN7;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trypsin I precursor (EC 3.4.21.4) (Cationic trypsinogen).
 GN PRSSI OR TRYLI OR TRPI OR TRYPI.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86221712; PubMed=30111602;
 RA Emi M., Nakamura Y., Ogawa M., Yamamoto T., Nishide T., Mori T.,
 RA Matsubara K.;
 RT "Cloning, Characterization and nucleotide sequences of two cDNAs
 RL encoding human pancreatic trypsinogens.";
 RL Gene 41:305-310(1986).
 RN [2]
 RP SEQUENCE OF 16-43.
 RX MEDLINE=90091010; PubMed=2598466;
 RA Kinland M., Russick C., Marks W.H., Borgstrom A.;
 RT "Immunoreactive anionic and cationic trypsin in human serum.";
 RL Clin. Chim. Acta 184:31-46(1989).
 RN [3]
 RP SEQUENCE OF 68-151 FROM N.A., AND VARIANT HPC HIS-122.
 RX MEDLINE=96438847; PubMed=8841182;
 RA Whitcomb D.C., Gorry M.C., Preston R.A., Furey W., Sossenheimer M.J.,
 RA Ulrich C.D., Martin S.P., Gates L.K. Jr., Amann S.T., Toskes P.P.,
 RA Liddle R., McGrath K., Uomo G., Post J.C., Ehrlich G.D.;
 RT "Hereditary pancreatitis is caused by a mutation in the cationic
 RL trypsinogen gene.";
 RL Nat. Genet. 14:141-145(1996).
 RN [4]
 RP SEQUENCE OF 15-67 FROM N.A., AND VARIANT HPC GLY-22.
 RX MEDLINE=20389982; PubMed=10930381;
 RA Telch N., Ockenga J., Hoffmeister A., Manns M., Mossner J., Keim V.;
 RT "Chronic pancreatitis associated with an activation peptide mutation
 RL that facilitates trypsin activation.";
 RL Gastroenterology 119:461-465(2000).
 RN [5]
 RP SEQUENCE OF 68-151 FROM N.A., AND VARIANTS HPC P-104; C-116 AND F-139.
 RA Telch N., Bauer N., Mossner J., Keim V.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS), AND PHOSPHORYLATION.
 RX MEDLINE=96266496; PubMed=8683601;
 RA Gaboriaud C., Serre L., Guy-Crotte O., Forest E.,
 RA Fontecilla-Camps J.-C.;
 RT "Crystal structure of human trypsin 1: unexpected phosphorylation of
 RL Tyr151.";
 RL J. Mol. Biol. 259:995-1010(1996).
 RN [7]
 RP VARIANTS HPC ILE-29 AND HIS-122.
 RX MEDLINE=97463797; PubMed=9322498;
 RA Gorry M.C., Gabaizadeh D., Furey W., Gates L.K. Jr., Preston R.A.,
 RA Aston C.E., Zhang Y., Ulrich C., Ehrlich G.D., Whitcomb D.C.;
 RT "Mutations in the cationic trypsinogen gene are associated with
 RL recurrent acute and chronic pancreatitis.";
 RL Gastroenterology 113:1063-1068(1997).
 RN [8]
 RP VARIANTS HPC ILE-29.
 RX MEDLINE=98295575; PubMed=9633818;
 RA Telch N., Mossner J., Keim V.;
 RT "Mutations of the cationic trypsinogen in hereditary pancreatitis.";
 RL Hum. Mutat. 12:39-43(1998).
 RN [9]
 RP VARIANTS HPC VAL-16 AND HIS-122.
 RX MEDLINE=99315544; PubMed=10381903;
 RA Witt H., Luck W., Becker M.;
 RT "A signal peptide associated with chronic pancreatitis.";
 RL The gene is strongly pleiotropic with chronic pancreatitis."

RL Biochim. Biophys. Acta 69:115-129(1963).
 RN [2]
 RP SEQUENCE OF 9-231.
 RX MEDLINE-73258692; PubMed-4738933;
 RA Hermodson M.A., Ericsson L.H., Neurath H., Walsh K.A.;
 RT "Determination of the amino acid sequence of porcine trypsin by
 RT sequenator analysis.";
 RL Biochemistry 12:3146-3153(1973).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
 RX MEDLINE-93187998; PubMed-8445634;
 RA Huang Q., Liu S., Tang Y.;
 RT "Refined 1.6-A resolution crystal structure of the complex formed
 RT between porcine beta-trypsin and MCTI-A, a trypsin inhibitor of the
 RT squash family. Detailed comparison with bovine beta-trypsin and its
 RT complex.";
 RL J. Mol. Biol. 229:1022-1030(1993).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
 RX MEDLINE-92201369; PubMed-1551419;
 RA Huang Q., Liu S., Tang Y., Zeng F., Qian R.;
 RT "Amino acid sequencing of a trypsin inhibitor by refined 1.6 A x-ray
 RT crystal structure of its complex with porcine beta-trypsin.";
 RL FEBS Lett. 297:143-146(1992).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
 RX MEDLINE-95035057; PubMed-7947985;
 RA Huang Q., Wang Z., Li Y., Liu S., Tang Y.;
 RT "Refined 1.8-A resolution crystal structure of the porcine epsilon-
 RT trypsin.";
 RL Biochim. Biophys. Acta 1209:77-82(1994).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF COMPLEX WITH LDTI.
 RX MEDLINE-97390427; PubMed-9242660;
 RA Stubbs M.T., Morenweiser R., Sturzebecher J., Bauer M., Bode W.,
 RA Huber R., Piechotta G.P., Matschner G., Sommerhoff C.P., Fritz H.,
 RA Auerwald E.A.;
 RT "The three-dimensional structure of recombinant leech-derived
 RT trypsin inhibitor in complex with trypsin. Implications for the
 RT structure of human mast cell trypsinase and its inhibition.";
 RL J. Biol. Chem. 272:19931-19937(1997).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.03 ANGSTROMS) OF COMPLEX WITH LDTI.
 RX MEDLINE-98046095; PubMed-9384562;
 RA di Marco S., Priestle J.P.;
 RT "Structure of the complex of leech-derived trypsinase inhibitor (LDTI)
 RT with trypsin and modeling of the LDTI-trypsinase system.";
 RL Structure 5:1465-1474(1997).
 CC -I- CATALYTIC ACTIVITY: Preferential cleavage: Arg-I-Xaa, Lys-I-Xaa.
 CC -I- SUBCELLULAR LOCATION: Extracellular.
 CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 DR PIR; A00947; TRPGF.
 DR PDB; IMCT; 31-JAN-94.
 DR PDB; LAKS; 12-FEB-97.
 DR PDB; LEPT; 07-FEB-95.
 DR PDB; LTFX; 21-JAN-98.
 DR PDB; LDTI; 20-MAY-98.
 DR PDB; LANI; 01-JUL-98.
 DR PDB; LAVN; 18-NOV-98.
 DR PDB; LAVX; 18-NOV-98.
 DR MEROPS; S01.151;
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen;
 KW 3D-structure.
 FT PROPEP 1 8 ACTIVATION PEPTIDE.
 FT CHAIN 9 231 TRYPSIN.

FT ACT_SITE 48
 FT ACT_SITE 92
 FT ACT_SITE 185
 FT DISULFID 15
 FT DISULFID 33
 FT DISULFID 49
 FT DISULFID 117
 FT DISULFID 124
 FT DISULFID 156
 FT DISULFID 181
 FT SITE 179
 FT VARIAT 20
 FT STRAND 10
 FT STRAND 13
 FT STRAND 14
 FT TURN 17
 FT TURN 21
 FT STRAND 23
 FT STRAND 31
 FT TURN 39
 FT STRAND 40
 FT STRAND 42
 FT HELIX 47
 FT STRAND 55
 FT STRAND 62
 FT TURN 63
 FT STRAND 71
 FT TURN 82
 FT TURN 83
 FT TURN 86
 FT TURN 90
 FT STRAND 94
 FT STRAND 112
 FT TURN 120
 FT STRAND 123
 FT STRAND 142
 FT STRAND 144
 FT STRAND 150
 FT HELIX 153
 FT TURN 161
 FT TURN 166
 FT STRAND 168
 FT TURN 174
 FT STRAND 179
 FT TURN 182
 FT TURN 185
 FT STRAND 188
 FT TURN 192
 FT STRAND 194
 FT STRAND 207
 FT TURN 208
 FT STRAND 210
 FT STRAND 212
 FT HELIX 217
 FT HELIX 221
 SQ SEQUENCE 231 AA; 24409 MW; A0A125CF7FC138C2_CRC64;
 Query Match 80.7%; Score 1071; DB 1; Length 231;
 Best Local Similarity 83.0%; Pred. No. 3.6e-87;
 Matches 191; Conservative 21; Mismatches 18; Indels 0; Gaps 0;
 Qy 17 PTDDDDKIVGYTCENSVPTQVSLNAGYHFCGGSLISDQVTVSAAHCKYSRIQVRLGEY 76
 Db 2 PTDDDDKIVGYTCENSVPTQVSLNAGYHFCGGSLISDQVTVSAAHCKYSRIQVRLGEH 61
 Qy 77 NIDVLEGNQFINSKAVIRHPNYSWILDNDMLIKLSSPAVLNARVATISLPRAAAG 136
 Db 62 NIDVLEGNQFINSKAVIRHPNYSWILDNDMLIKLSSPAVLNARVATISLPRAAAG 121
 Qy 137 TQCLISGWNTLSSGTNYPELLQCLDAPILTAQACEASYPGQITENMICAGFLEGGKDCS 196
 Db 122 TECLISGWNTKSSGSYPYSLQLCLKAPVLSDDSCCKSSYPGQITGNMICVGFLEGGKDCS 181
 Qy 197 QGDSGGPVVNCNLOGIVSWGYGCAQKNKPGYTVKVCNFWDIQSTIAAN 246
 Db 182 QGDSGGPVVNCNLOGIVSWGYGCAQKNKPGYTVKVCNFWDIQSTIAAN 231

CHARGE RELAY SYSTEM.
 CHARGE RELAY SYSTEM.
 CHARGE RELAY SYSTEM.

REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 I -> V.

RESULT 10

```

TRY3_HUMAN
ID TRY3_HUMAN STANDARD; PRT; 247 AA.
AC P15951;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin III precursor (EC 3.4.21.4).
GN PRSS3 OR TRY3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
TISSUE=Pancreas;
MEDLINE=90221895; PubMed=2326201;
Tanai T., Kawashima I., Mita K., Takiguchi Y.;
"Nucleotide sequence of the human pancreatic trypsinogen III cDNA.";
Nucleic Acids Res. 18:1631-1631(1990).
-!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-I-Xaa, Lys-I-Xaa.
-!- SUBCELLULAR LOCATION: Extracellular.
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X15505; CAA33527.1; -
CC PIR; S12764; S12764.
CC HSP; P00761; IEPT.
CC MEROPS; S01.174; -.
CC Genew; HGNC:9486; PRSS3.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00020; TRYPSIN_DOM; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; FALSE_NEG.
CC Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
CC Multigene family.
CC SIGNAL 1 15
CC PROPEP 16 23 ACTIVATION PEPTIDE.
CC CHAIN 24 247 TRYPSIN III.
CC ACT_SITE 63 63 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT_SITE 107 107 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT_SITE 200 200 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT_SITE 30 160 BY SIMILARITY.
CC DISULFID 48 64 BY SIMILARITY.
CC DISULFID 139 206 BY SIMILARITY.
CC DISULFID 171 185 BY SIMILARITY.
CC SITE 194 194 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
CC SEQUENCE 247 AA; 26776 MW; 697DE163FCE0D6 CRC64;
CC
CC Query Match 79.7% Score 1058; DB 1; Length 247;
CC Best Local Similarity 78.5% Pred. No. 5.4e-86;
CC Matches 194; Conservative 27; Mismatches 26; Indels 0; Gaps 0;
QY 1 MNPLLIALFAGAAVATDDDDKIVGGYTCENSVQVSLNAGYHFGGSLIDQWVVS 60
DB 1 MNPLLIALFAGAAVATPDDDDKIVGGYTCENSVQVSLNAGYHFGGSLIDQWVVS 60
QY 61 AAHCYKSRIOVRGGEYNDVLEGEQFINSKVIHRPNYSWILNDIMLIKSSPAVLN 120
DB 61 AAHCYKTRIOVRGGEYNDVLEGEQFINSKVIHRPNYSWILNDIMLIKSSPAVLN 120
QY 121 ARVATISLPACRAAGTQCLISGNGNTLSGGTNPPELLQCLDAPILTAQCEASYPGOIT 180

```

RESULT 11

```

TRY4_HUMAN
ID TRY4_HUMAN STANDARD; PRT; 304 AA.
AC P35030; Q15665; Q9UQW3;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin IV precursor (EC 3.4.21.4) (Brain trypsinogen)
DE (Mesotrypsinogen).
GN PRSS4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORMS A AND B).
TISSUE=Brain;
MEDLINE=94123994; PubMed=8294000;
Wiegand U., Corbach S., Minn A., Kang J., Mueller-Hill B.;
"Cloning of the cDNA encoding human brain trypsinogen and
characterization of its product.";
Gene 136:167-175(1993).
RN [2]
SEQUENCE FROM N.A. (ISOFORM C).
RA Fukuoka S.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-I-Xaa, Lys-I-Xaa.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- ALTERNATIVE PRODUCTS: 3 isoforms; A (shown here), B and C; seem
CC to be produced by alternative splicing.
CC -!- TISSUE SPECIFICITY: PANCREAS AND BRAIN.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X72781; CAB58178.1; -
CC EMBL; X71345; CAA50484.1; -
CC EMBL; D45417; BAA08257.1; -
CC PIR; S33496; S33496.
CC HSP; P07477; ITRN.
CC MEROPS; S01.174; -.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00020; TRYPSIN_DOM; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; FALSE_NEG.
CC Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
CC Multigene family; Alternative splicing.
CC SIGNAL 1 80 POTENTIAL.
CC PROPEP 81 304 ACTIVATION PEPTIDE.
CC CHAIN 120 120 TRYPSIN IV.
CC ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).

```


CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M11590; AAA30900.1; -
 DR PIR; B26273; TRDGC.
 DR HSSP; P00761; LEPT.
 DR MEROPS; S01.151; -
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS02040; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
 KW Multigene family.
 T SIGNAL 1 15
 I PROPEP 16 23 ACTIVATION PEPTIDE.
 FT CHAIN 24 246 TRYPSIN, CATIONIC.
 FT ACT_SITE 63 63 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 107 107 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 200 200 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 30 160 BY SIMILARITY.
 FT DISULFID 48 64 BY SIMILARITY.
 FT DISULFID 132 233 BY SIMILARITY.
 FT DISULFID 139 206 BY SIMILARITY.
 FT DISULFID 171 185 BY SIMILARITY.
 FT DISULFID 196 220 BY SIMILARITY.
 FT SITE 194 194 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 SQ SEQUENCE 246 AA; 26170 MW; E9E5ALDE391BBB CRC64;

Query Match 77.3%; Score 1026; DB 1; Length 246;
 Best Local Similarity 75.2%; Pred. No. 3.5e-83;
 Matches 185; Conservative 23; Mismatches 38; Indels 0; Gaps 0;

QY 1 MNPLLIAGNAVATPDDDDKIVGGYTCSENSVYQVSLNAGYHFCGSLSDQWVYS 60
 Db 1 MKTFILALGATVAFPIDDDDKIVGGYTCSENSVYQVSLNAGYHFCGSLSDQWVYS 60
 QY 61 AAHCYKSRIOVRLGEYNDVLEGNQFINSKVRHPNYSWILNDIMLKLSFPAVLN 120
 Db 61 AAHCYKSRIOVRLGEYNDVLEGNQFINSKVRHPNYSWILNDIMLKLSFPAVLN 120
 QY 121 ARVATISLPACAAAGTQCLISGNGNTLSGNTYPELLOCLDAPILTOAOCEASYPGOIT 180
 Db 121 SRVSAIALPKSCPAAGTQCLISGNGNTQSGONTPDVLQCLKAPILSDVCNATPGQIS 180
 QY 181 ENMTAGFLGGKDCSGDGGPVVCGELQGIYSWGWGCAQKNKPKGYTTKVCNFDVWQ 240
 Db 181 SNMCLGYMEGKDCSGDGGPVVCGELQGIYSWGWGCAQKNKPKGYTTKVCNFDVWQ 240
 QY 241 STIAAN 246
 Db 241 QTIAAN 246

RESULT 14
 TRYL_BOVIN
 ID TRYL_BOVIN STANDARD; PRT; 243 AA.
 AC P00760;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trypsinogen, cationic precursor (EC 3.4.21.4) (Beta-trypsin)
 DE (Fragment).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;

RN [1] SEQUENCE FROM N.A.
 RP TISSUE-Pancreas;
 RC Okajima T., Maniwa M., Nagao S., Fujikawa H., Goto S.;
 RA Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 15-243, AND DISULFIDE BONDS.
 RX MEDLINE-67168848; PubMed-5967094;
 RA Mikes O., Holeysovsky V., Tomasek V., Sorm F.;
 RT "Covalent structure of bovine trypsinogen. The position of the
 RT remaining amides";
 RL Biochem. Biophys. Res. Commun. 24:346-352(1966).
 RN [3]
 RP REVISIONS.
 RX MEDLINE-72035053; PubMed-4399051;
 RA Hartley B.S.;
 RT "Homologies in serine proteinases";
 RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 257:77-87(1970).
 RN [4]
 RP REVISIONS.
 RX MEDLINE-75146445; PubMed-1092332;
 RA Titani K., Ericsson L.H., Neurath H., Walsh K.A.;
 RT "Amino acid sequence of dogfish trypsin";
 RL Biochemistry 14:1358-1366(1975).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF CALCIUM-BINDING SITE.
 RX MEDLINE-76072097; PubMed-512;
 RA Bode W., Schwager P.;
 RT "The refined crystal structure of bovine beta-trypsin at 1.8-A
 RT resolution. II. Crystallographic refinement, calcium binding site,
 RT benzamide binding site and active site at pH 7.0.";
 RL J. Mol. Biol. 98:693-717(1975).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
 RX MEDLINE-77112431; PubMed-556951;
 RA Kosiakoff A.A., Chambers J.L., Kay L.M., Stroud R.M.;
 RT "Structure of bovine trypsinogen at 1.9-A resolution.";
 RL Biochemistry 16:654-664(1977).
 RN [7]
 RP DISULFIDE BONDS.
 RX MEDLINE-66079271; PubMed-5892911;
 RA Kauffman D.L.;
 RT "The disulphide bridges of trypsin";
 RL J. Mol. Biol. 12:929-932(1965).
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN THE ACINAR CELLS OF THE
 CC PANCREAS.
 CC -1- PTM: AUTOCATALYTIC CLEAVAGE AFTER LYS-20 LEADS TO BETA-TRYPSIN BY
 CC RELEASING A TERMINAL HEXAPEPTIDE. SUBSEQUENT CLEAVAGE AFTER
 CC LYS-145 LEADS TO ALPHA-TRYPSIN. FURTHER CLEAVAGE AFTER LYS-190
 CC YIELDS PSEUDOTRYPSIN. A CLEAVAGE MAY ALSO OCCUR AFTER ARG-119.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- DATABASE: NAME=worthington enzyme manual;
 CC WWW="http://www.worthington-biochem.com/manual/TRY.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D38507; BAA07516.1; -
 DR PIR; A00946; TRBOTR.
 DR PDB; 2TGA; 09-APR-85.
 DR PDB; 1TGB; 14-MAR-85.
 DR PDB; 1TGC; 09-APR-85.
 DR PDB; 2TGD; 07-MAY-86.
 DR PDB; 1TGN; 22-OCT-84.
 DR PDB; 2TGP; 14-MAR-85.
 DR PDB; 1TGS; 14-MAR-85.

```
DR PDB; 1TGT; 15-OCT-90.
DR PDB; 2TGT; 09-APR-85.
DR PDB; 1TLD; 15-OCT-92.
DR PDB; 2TLD; 15-JUL-92.
DR PDB; 1TPA; 14-MAR-85.
DR PDB; 2TPI; 15-JUL-91.
DR PDB; 3TPI; 14-MAR-85.
DR PDB; 4TPI; 08-NOV-85.
DR PDB; 1TPO; 14-MAR-85.
DR PDB; 1TPP; 16-APR-87.
DR PDB; 1TPB; 15-APR-91.
DR PDB; 3TPB; 14-MAR-85.
DR PDB; 2ETC; 14-MAR-85.
DR PDB; 2PTN; 15-APR-91.
DR PDB; 3PTN; 09-APR-85.
DR PDB; 5PTP; 07-JUL-97.
DR PDB; 1TAB; 15-JUL-92.
DR PDB; 1GBT; 31-JAN-94.
DR PDB; 1PPE; 31-JAN-94.
DR PDB; 1PPE; 31-JAN-94.
DR PDB; 1PPH; 31-JAN-94.
DR PDB; 1SMF; 31-JUL-94.
DR PDB; 1TNG; 30-NOV-94.
DR PDB; 1TNH; 30-NOV-94.
DR PDB; 1TNI; 30-NOV-94.
DR PDB; 1TNJ; 30-NOV-94.
DR PDB; 1TNK; 30-NOV-94.
DR PDB; 1TNL; 30-NOV-94.
DR PDB; 1BTP; 29-JAN-96.
DR PDB; 1BTW; 15-OCT-95.
DR PDB; 1BTX; 15-OCT-95.
DR PDB; 1BTY; 15-OCT-95.
DR PDB; 1BTZ; 15-OCT-95.
DR PDB; 1JRS; 14-OCT-96.
DR PDB; 1JRT; 14-OCT-96.
DR PDB; 1MAX; 14-OCT-96.
DR PDB; 1MAX; 14-OCT-96.
DR PDB; 1MAY; 14-OCT-96.
DR PDB; 1MTS; 20-AUG-97.
DR PDB; 1MTU; 12-NOV-97.
DR PDB; 1MTV; 12-NOV-97.
DR PDB; 1MTW; 12-NOV-97.
DR PDB; 1TPS; 26-JAN-95.
DR PDB; 1TYN; 26-JAN-95.
DR PDB; 1TAW; 24-JUN-97.
DR PDB; 1AQ7; 25-FEB-98.
DR PDB; 1TIO; 23-SEP-98.
DR PDB; 2TIO; 30-SEP-98.
DR PDB; 1XUG; 16-DEC-98.
DR PDB; 1XUG; 16-DEC-98.
DR PDB; 1XUH; 11-NOV-98.
DR PDB; 1XUI; 11-NOV-98.
DR PDB; 1XUJ; 11-NOV-98.
DR PDB; 1XUK; 11-NOV-98.
DR PDB; 1AUJ; 14-OCT-98.
DR PDB; 1AZ8; 13-JAN-99.
DR PDB; 1BJU; 13-JAN-99.
DR PDB; 1BJV; 13-JAN-99.
DR PDB; 2BZA; 23-MAR-99.
DR PDB; 1CE5; 23-MAR-99.
DR MEROPS; S01.151; -.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR SMART; SM00020; Tryp_Src; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
3D-structure.
FT NON_TER 1
FT SIGNAL <1 14
FT PROPEP 15 20 ACTIVATION PEPTIDE.
FT CHAIN 21 243 TRYPsin, CATIONIC.
FT CHAIN 21 145 ALPHA-TRYPSIN CHAIN 1.
```

```
FT CHAIN 146 243 ALPHA-TRYPSIN CHAIN 2.
FT CA_BIND 72 82 SUBSTRATE.
FT BINDING 191 192 SUBSTRATE.
FT BINDING 194 195 SUBSTRATE.
FT BINDING 197 197 SUBSTRATE.
FT DISULFID 27 157
FT DISULFID 45 61
FT DISULFID 129 230
FT DISULFID 136 203
FT DISULFID 168 182
FT DISULFID 193 217
FT ACT_SITE 60 60 CHARGE RELAY SYSTEM.
FT ACT_SITE 104 104 CHARGE RELAY SYSTEM.
FT ACT_SITE 197 197 CHARGE RELAY SYSTEM.
FT SITE 191 191 REQUIRED FOR SPECIFICITY.
FT STRAND 22 22
FT STRAND 25 26
FT TURN 29 30
FT TURN 33 34
FT STRAND 35 39
FT STRAND 43 51
FT TURN 52 53
FT STRAND 54 57
FT STRAND 59 61
FT STRAND 67 70
FT TURN 75 76
FT STRAND 83 92
FT TURN 94 95
FT STRAND 97 97
FT TURN 98 101
FT STRAND 102 102
FT TURN 103 103
FT STRAND 106 110
FT STRAND 124 124
FT TURN 132 133
FT STRAND 135 140
FT STRAND 156 162
FT STRAND 165 171
FT TURN 173 174
FT TURN 178 179
FT STRAND 180 183
FT TURN 186 187

Query Match 76.8%; Score 1019; DB 1; Length 243;
Best Local Similarity 76.4%; Pred. No. 1.4e-82;
Matches 185; Conservative 30; Mismatches 27; Indels 0; Gaps 0;

QY 5 LILAFLLGAATPTDDDDKIVGGYTCCEENSVPYQVSLNAGYHFCGSLISDQWVYVSAHC 64
Db : ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
2 IFLLALGAATVPPVDDDDKIVGGYTCGANTVPYQVSLNAGYHFCGSLISQWVYVSAHC 61
QY 65 YKSRIQVRLGEYNIDVLENEQFINSKVIIRHPNYSWILNDIMLIKLSPPAVLNARVA 124
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
62 YKSGIQVRLGEDINIVVEGNEQFISAKSVIHPNSYNTLNNDIMLIKLSAASLSRVA 121
QY 125 TISLPACAPGTQCLISGNGNTLSSTNYPVPELLOCLDAPILTQACEASYPGQITENMI 184
Db : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
122 SISLPTSCASAGTQCLISGNGNTKSGTSYPDVILKCLKAPILSDSCSKSAYPGQITSNMF 181
QY 185 CAGFLEGGKDSQGDGGPVCNGELQGVSWCYGCAQKNKPGVYTKVCNFVDWIQSTIA 244
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
182 CAGYLEGGKDSQGDGGPVCNGELQGVSWCYGCAQKNKPGVYTKVCNFVDWIQSTIA 241
QY 245 AN 246
Db :
242 SN 243

RESULT 15
TRY1_XENLA
ID TRY1_XENLA STANDARD; PRT; 243 AA.
AC P19799;
DT 01-FEB-1991 (Rel. 17, Created)
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 15:14:41 ; Search time 79 Seconds

(without alignments)
644.223 Million cell updates/sec

Title: US-09-762-277A-1

Perfect score: 1327

Sequence: 1 MNPLILIAFLGAATPTDD.....VYTKVNFVDWQSTIAANS 247

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_pbc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1119	84.3	246	11 Q90UK9	Q90UK9 mus musculus
2	1118	84.3	246	11 Q90T07	Q90T07 mus musculus
3	1099	82.8	246	11 Q921R9	Q921R9 mus musculus
4	1063.5	80.1	247	11 Q9CFN9	Q9CFN9 mus musculus
5	1056.5	79.6	247	11 Q9D7Y7	Q9D7Y7 mus musculus
6	970.5	73.1	247	11 Q9CFN7	Q9CFN7 mus musculus
7	906	68.3	241	13 Q987G9	Q987G9 engraulis j
8	899.5	67.8	237	13 Q91515	Q91515 fugu rubrip
9	897	67.6	242	13 Q93266	Q93266 pseudopleur
10	887	66.8	240	13 Q987H0	Q987H0 pseudopleur
11	879	66.2	242	13 Q9W7Q7	Q9W7Q7 paralicthy
12	872	65.7	244	13 Q80CW3	Q80CW3 anguilla ja
13	871.5	65.7	238	13 Q9W7G6	Q9W7G6 paralicthy
14	854	64.4	242	13 Q92099	Q92099 paranotothe
15	838.5	63.2	247	13 Q42608	Q42608 petromyzon
16	837.5	63.1	247	13 Q42158	Q42158 petromyzon

17	834	62.8	245	13	Q42160	petromyzon
18	828	62.4	244	13	Q42159	petromyzon
19	773.5	58.3	219	13	Q91036	gadus morhu
20	765	57.6	247	13	Q9W7Q5	paralicthy
21	730.5	55.0	250	13	Q93265	pseudopleur
22	699	52.7	249	13	Q92046	dissostichu
23	698	52.6	249	13	Q9W6K0	notothenia
24	679.5	51.2	178	13	Q93594	dicentrarch
25	679.5	51.2	675	13	Q9W6J8	dissostichu
26	658.5	49.6	344	13	Q9W6J9	dissostichu
27	633	47.7	138	11	Q9D7Z5	mus musculu
28	586.5	44.2	293	11	Q9D140	mus musculu
29	570.5	43.0	255	4	Q96RQ0	mus sapien
30	564	42.5	146	13	Q9DDE1	brachydanio
31	560.5	42.2	249	11	Q9QYN4	mus musculu
32	560.5	42.2	276	11	Q9QYN3	mus musculu
33	554.5	41.8	267	5	Q9BK47	luidia foli
34	552	41.6	263	5	Q62562	penaeus van
35	551	41.5	251	11	Q54854	rattus norv
36	541	40.8	246	11	Q88301	mus musculu
37	541	40.8	253	11	Q91Y82	mus musculu
38	537.5	40.5	234	11	Q9CV76	mus musculu
39	537	40.5	264	5	Q62561	penaeus van
40	534	40.2	266	5	Q27761	penaeus van
41	531	40.0	262	4	Q8TCV8	homo sapien
42	530.5	40.0	261	6	Q29474	canis famil
43	530	39.9	263	5	Q9TY16	penaeus van
44	528.5	39.8	249	11	Q91VE3	mus musculu
45	523.5	39.4	233	13	Q9PT51	agkistrodon

ALIGNMENTS

RESULT 1

Q90UK9 Q90UK9 PRELIMINARY; PRT; 246 AA.
AC Q90UK9;
DT 01-MAY-2000 (TREMREL. 13, Created)
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)
DT 01-MAR-2002 (TREMREL. 20, Last annotation update)
DE TESP4 (Q91000IB19RIK protein) (Trypsinogen 9).
GN TC OR Q91000IB19RIK OR TRYPSINOGEN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE=99436155; PubMed=10506205;
RA Ohmura K., Kohno N., Kobayashi Y., Yamagata K., Sato S.,
RA Kashiwabara S., Baba T.;
RT "A homologue of pancreatic trypsin is localized in the acrosome of
mammalian sperm and is released during acrosome reaction.";
RL J. Biol. Chem. 274:29426-29432(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kohno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldi M.F.,
RA Brownstein M.J., Buit C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsuke S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RX MEDLINE=21103195; PubMed=11160223;
 RA Chen F., Rowen L., Hood L., Rothenberg E.V.;
 RT "Differential transcriptional regulation of individual TCR Vbeta
 RT segments before gene rearrangement.";
 RL J. Immunol. 166:1771-1780(2001).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 TRYPsin FAMILY.
 DR EMBL; AB009661; BAA45187.1; -;
 DR EMBL; AB017031; BAA74760.1; -;
 DR EMBL; AK007843; BAB25300.1; -;
 DR EMBL; AE000664; AAB69057.1; -;
 DR HSSP; P00763; 1DPO.
 DR MEROPS; S01.058; -;
 DR MGD; MGI:1913350; 0910001B19Rik.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser.protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPsin.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 246 AA; 26277 MW; 915C92353EC42809 CRC64;

Query Match 84.3%; Score 1119; DB 11; Length 246;
 Best Local Similarity 81.1%; Pred. No. 1.2e-95;
 Matches 201; Conservative 22; Mismatches 23; Indels 0; Gaps 0;
 QY 1 MNPLILAFILGAATPDDDDKIYGVTCENSVYQVSLNAGYHFCGSLSDQWVVS 60
 DB 1 MNSLLFLALVGAATPDDDDKIYGVTCENSVYQVSLNAGYHFCGSLSDQWVVS 60
 QY 61 AAHCYKSRIOVRIGYNDVLEQNFINSKVIKHPNYSWILNDIMLIKSSPAVLN 120
 DB 61 AAHCYKSRIOVRIGYNDVLEQNFINSKVIKHPNYSWILNDIMLIKSSPAVLN 120
 QY 121 ARVATISLPACAPAGTQCLISGWGNTLSSTGNTYPELLQCLDAPILITQACEASYPGQIT 180
 DB 121 ARVATVALPSSCAPAGTQCLISGWGNTLSSTGNTYPELLQCLDAPILITQACEASYPGKIT 180
 QY 181 ENMICAGLEGKDCQSGDGGPVVNGELQIVSWGCGCAKPKGVYTKVCFVDTIQ 240
 DB 181 NMHCIVGLEKDCQSGDGGPVVNGELQIVSWGCGCAKPKGVYTKVCFVDTIQ 240
 QY 241 STIAAN 246
 DB 241 DTIAAN 246

RESULT 2

Q9ROT7 PRELIMINARY; PRT; 246 AA.
 AC Q9ROT7;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DE 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Paucitric trypsin (0910001B19Rik protein) (Trypsinogen 8).
 GN TD OR 0910001B19RIK OR TRYPSINOGEN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

RN SEQUENCE FROM N.A.
 RP STRAIN=129SVJ;
 RX MEDLINE=99436155; PubMed=10506205;
 RA Omura K., Kohno N., Kobayashi Y., Yamagata K., Sato S.,
 RA Kashiwabara S., Baba T.;
 RT "A homologue of pancreatic trypsin is localized in the acrosome of
 RT mammalian sperm and is released during acrosome reaction.";
 RL J. Biol. Chem. 274:29426-29432(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=STOMACH, AND SPLEEN;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsuke S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RX MEDLINE=21103195; PubMed=11160223;
 RA Chen F., Rowen L., Hood L., Rothenberg E.V.;
 RT "Differential transcriptional regulation of individual TCR Vbeta
 RT segments before gene rearrangement.";
 RL J. Immunol. 166:1771-1780(2001).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 TRYPsin FAMILY.
 DR EMBL; AB017032; BAA74761.1; -;
 DR EMBL; AK008667; BAB25821.1; -;
 DR EMBL; AK003064; BAB22542.1; -;
 DR EMBL; AE000664; AAB69056.1; -;
 DR HSSP; P00763; 1SLU.
 DR MEROPS; S01.057; -;
 DR MGD; MGI:1913350; 0910001B19Rik.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser.protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPsin.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 246 AA; 26274 MW; B6A9F4C99079633F CRC64;

Query Match 84.3%; Score 1118; DB 11; Length 246;
 Best Local Similarity 82.1%; Pred. No. 1.5e-95;
 Matches 202; Conservative 21; Mismatches 23; Indels 0; Gaps 0;
 QY 1 MNPLILAFILGAATPDDDDKIYGVTCENSVYQVSLNAGYHFCGSLSDQWVVS 60
 DB 1 MRALLFLALVGAATPDDDDKIYGVTCENSVYQVSLNAGYHFCGSLSDQWVVS 60
 QY 61 AAHCYKSRIOVRIGYNDVLEQNFINSKVIKHPNYSWILNDIMLIKSSPAVLN 120
 DB 61 AAHCYKSRIOVRIGYNDVLEQNFINSKVIKHPNYSWILNDIMLIKSSPAVLN 120

```

QY 121 ARVATISLPRACAAPGTCQLISGWNTLSSTNTYVELLOCLDAPILTAQACEASYPGKIT 180
DB 121 ARVATVALPSSCAPAGTCQLISGWNTLSFGVNNPDLLOCLDAPLLPADCEASYPGKIT 180
QY 181 ENMICAGLEGGKSCQDGGPVVNCNGELQIVSWGCGCAQKNKPGVYTKVCFNFDWIQ 240
DB 181 NMICVGLFEGGKSCQDGGPVVNCNGELQIVSWGCGCAQKNKPGVYTKVCFNFDWIQ 240
QY 241 STIAAN 246
DB 241 NTIAAN 246

RESULT 3
Q921R9 PRELIMINARY; PRT; 246 AA.
AC Q921R9
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
GN TRYPSINogen 16.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Balb/C;
RA Rowen L., Hood L.;
RT *Comparison between strains Balb/C and 129 in a region of the mouse T
RL cell receptor beta locus";
RN Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Balb/C;
RA Chen F., Rowen L., Hood L., Rothenberg E.V.;
RT "Differential transcriptional regulation of individual TCR Vbeta
RN segments before gene rearrangement.";
CC J. Immunol. 166:1771-1780(2001).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AF107342; AAC79093.1; -
DR EMBL; AF000655; AAB69088.1; -
DR HSSP; P00763; IDPO.
DR MEROPS; S01.063; -
DR MGD; MGI:2148749; Trygln6.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPsin.
DR SMART; SM00020; TRYp_SPc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 246 AA; 26134 MW; 34E173B18CA2F463 CRC64;

Query Match 82.8%; Score 1099; DB 11; Length 246;
Best Local Similarity 80.9%; Pred. No. 8.9e-94;
Matches 199; Conservative 23; Mismatches 24; Indels 0; Gaps 0;

QY 1 MNPILLALFAGAAVPTDDDDKIVGGYTCENSVYPYQVSLNAGYFCGSLINDQWVVS 60
DB 1 MSALLFLALVGAFAVFPDDDDKIVGGYTCENSVYPYQVSLNAGYFCGSLINDQWVVS 60
QY 61 AAHCYKSRIOVLGEINIDVLENEQFINSAKIVIRHPNYSWILDNDIMLIKLSPPAVLN 120
DB 61 AAHCYKTRIOVLGEININYLENEQFINSAKIVIRHPNYSWILDNDIMLIKLSPPAVLN 120
QY 121 ARVATISLPRACAAPGTCQLISGWNTLSSTNTYVELLOCLDAPILTAQACEASYPGKIT 180
DB 121 ARVATVALPSSCAPAGTCQLISGWNTLSFGVNNPDLLOCLDAPLLPADCEASYPGKIT 180

```

```

DB 121 ARVATVALPSSCAPAGTCQLISGWNTLSFGVSEPDLLQCLDAPLLPADCEASYPGKIT 180
QY 181 ENMICAGLEGGKSCQDGGPVVNCNGELQIVSWGCGCAQKNKPGVYTKVCFNFDWIQ 240
DB 181 NMICVGLFEGGKSCQDGGPVVNCNGELQIVSWGCGCAQKNKPGVYTKVCFNFDWIQ 240
QY 241 STIAAN 246
DB 241 DTIAAN 246

RESULT 4
Q9CPN9 PRELIMINARY; PRT; 247 AA.
AC Q9CPN9
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE 2210010C04RIK protein (Trypsinogen 7).
GN 2210010C04RIK OR TRYPSINOGEN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=SMALL INTESTINE, AND PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Balb/C;
RX MEDLINE=21103195; PubMed=11160223;
RA Chen F., Rowen L., Hood L., Rothenberg E.V.;
RT "Differential transcriptional regulation of individual TCR Vbeta
RN segments before gene rearrangement.";
CC J. Immunol. 166:1771-1780(2001).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AK008460; BAB25680.1; -
DR EMBL; AK007773; BAB25246.1; -
DR EMBL; AE000663; AAB69044.1; -
DR HSSP; P00763; IDPO.
DR MEROPS; S01.151; -
DR MGD; MGI:1914623; 2210010C04RIK.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPsin.
DR SMART; SM00020; TRYp_SPc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
KW

```


SO SEQUENCE 247 AA; 26422 MW; B8C5767B18D9AD CRC64;

Query Match 80.1%; Score 1063.5; DB 11; Length 247;
Best Local Similarity 76.1%; Pred. No. 1.7e-90;

Matches 188; Conservative 31; Mismatches 27; Indels 1; Gaps 1;

QY 1 MNPILLIPLAFLGAAPV-TDDDDKIYGGYTCSEENSVYQVSLNAGYHFCGSLISQWV 59
DB 1 MKTLIFLAFLGAVALPLDDDDKIYGGYTCORNALPYQVSLNAGYHFCGSLISQWV 60
QY 60 SAACHYKSRIOVRLGEYNDVLEGNQFINSKVIIRHPNYSWILNDIMLIKSSPAVL 119
DB 61 SAACHYKSRIOVRLGEYNDVLEGNQFINSKVIIRHPNYSWILNDIMLIKSSPAVL 120
QY 120 NARVATISLPACAPGTCTCLISGNGNTLSGNTNYPQLQCLDAPILTOAOCEASYPGOI 179
DB 121 NSRVSTVALPRSCPSAGTCLVSGWNTLSGNTNYPQLQCLDAPILTOAOCEASYPGOI 180
QY 180 TENMTCAGLEGKSCQSGDGGPVVNCNGELQIVSWGCGCAOKNPGVYTKVCFVDWI 239
DB 181 TSNMFCGLGLEGKSCQSGDGGPVVNCNGELQIVSWGCGCAOKNPGVYTKVCFVDWI 240
QY 240 QSTIAAN 246
DB 241 QOTIAAN 247

RESULT 5

Q9D7Y7 PRELIMINARY; PRT; 247 AA.
AC Q9D7Y7
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 2210010C04Rik protein.
GN 2210010C04Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=STOMACH;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momabets P.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE

CC TRYPSIN FAMILY.

DR EMBL; AK008695; BAB25837.1; -

DR HSSP; P00763; 1DPO.

DR MEROPS; S01.151; -

DR MGD; MGI:1914623; 2210010C04Rik.

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR001254; Ser.protease_Try.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM0020; TRYD-SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS0134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS0135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 247 AA; 26407 MW; 84D474DB322B3A55 CRC64;

Query Match 79.6%; Score 1056.5; DB 11; Length 247;
Best Local Similarity 75.7%; Pred. No. 7.7e-90;
Matches 187; Conservative 31; Mismatches 28; Indels 1; Gaps 1;

QY 1 MNPILLIPLAFLGAAPV-TDDDDKIYGGYTCSEENSVYQVSLNAGYHFCGSLISQWV 59
DB 1 MKTLIFLAFLGAVALPPDDDDKIYGGYTCORNALPYQVSLNAGYHFCGSLISQWV 60
QY 60 SAACHYKSRIOVRLGEYNDVLEGNQFINSKVIIRHPNYSWILNDIMLIKSSPAVL 119
DB 61 SAACHYKSRIOVRLGEYNDVLEGNQFINSKVIIRHPNYSWILNDIMLIKSSPAVL 120
QY 120 NARVATISLPACAPGTCTCLISGNGNTLSGNTNYPQLQCLDAPILTOAOCEASYPGOI 179
DB 121 NSRVSTVALPRSCPSAGTCLVSGWNTLSGNTNYPQLQCLDAPILTOAOCEASYPGOI 180
QY 180 TENMTCAGLEGKSCQSGDGGPVVNCNGELQIVSWGCGCAOKNPGVYTKVCFVDWI 239
DB 181 TSNMFCGLGLEGKSCQSGDGGPVVNCNGELQIVSWGCGCAOKNPGVYTKVCFVDWI 240
QY 240 QSTIAAN 246
DB 241 QOTIAAN 247

RESULT 6

Q9CPN7 PRELIMINARY; PRT; 247 AA.
AC Q9CPN7
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 1810009J06Rik protein (Trypsinogen 4).
GN 1810009J06Rik OR TRYPSINOGEN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momabets P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;

```

RX MEDLINE=21103195; PubMed=11160223;
RA Chen F., Rowen L., Hood L., Rothenberg E.V.;
RT "Differential transcriptional regulation of individual TCR Vbeta
RL segments before gene rearrangement.";
RL J. Immunol. 166.1771-1780(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
DR EMBL; AK007406; BAB25018.1; -.
DR EMBL; AE000663; AAB69055.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.129; -.
DR MGD; MGI:1920876; 1810009J06Rik.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolyase; Serine protease.
DR SEQUENCE 247 AA; 26503 MW; EDF5F7696833C7BC CRC64;

Query Match 73.1%; Score 970.5; DB 11; Length 247;
Best Local Similarity 70.4%; Pred. No. 7e-82;
Matches 174; Conservative 31; Mismatches 41; Indels 1; Gaps 1;

QY 1 MNPLLIALLFAGAAVATPTDDDDKIVGGYTCENSVYQVSLNAGY-HFCGSLSDQWV 59
Db 1 MKIITFTFLGAALVALPANSDDKIVGGYTCPKHSVPYQVSLNDGISHQCGSLSDQWV 60
QY 60 SAAHCYKSRIOVRLGEYNIDVLEGNQFINSKAVIRHPNYSWILNDIMLIKLSPAVL 119
Db 61 SAAHCYKRRIOVRLGEYNIDVLEGNQFIDAELIIRHPDYNKDTVDNDIMLIKLSPAVL 120
QY 120 NARVATISLPRAAAPGTCQLISGWNTLSSGTNYPELLQCLDAPILTAQCEASYPGQI 179
Db 121 NSQVSTVSLPRSCASTNAQCLVSGWNTVSIIGGYPALLOCLEAPVLSASSCKSYPGQI 180
QY 180 TENNICAGFLEGGKDCQSGSGPVVNCNGELQIVSGYGAQKNGPVYTKVCFNFDWI 239
Db 181 TSNMFLGLFEGGKDCQSGSGPVVNCNGELQIVSGYGAQKNGPVYTKVCFNFDWI 240
QY 240 QSTIAAN 246
Db 241 QETWANN 247

RESULT 7
ID Q98TG9 PRELIMINARY; PRT; 241 AA.
AC Q98TG9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Trypsinogen II.
GN ATRYII.
OS Engraulis japonicus (Japanese anchovy).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Engraulidae;
OC Engraulis.
OX NCBI_TaxID=42892;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PYLORIC CAECA;
RA Watabe S., Ahsan M.N., Funabara D.;
RT "Anchovy trypsinogen mRNA.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
DR EMBL; AB041930; BAB40330.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.258; -.

```

```

DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolyase; Serine protease.
DR SEQUENCE 241 AA; 26282 MW; FE362D39CAEB2F6 CRC64;

Query Match 68.3%; Score 906; DB 13; Length 241;
Best Local Similarity 67.2%; Pred. No. 6.4e-76;
Matches 164; Conservative 34; Mismatches 40; Indels 6; Gaps 2;

QY 1 MNPLLIALLFAGAAVATPTDDDDKIVGGYTCENSVYQVSLNAGYHFCGSLSDQWV 60
Db 1 MRSILVFLVLLGAAFA---EDDKIVGGYTCQPSQPHQVSLNSGVHFCGSLSDQWV 56
QY 61 AACHYKSRIOVRLGEYNIDVLEGNQFINSKAVIRHPNYSWILNDIMLIKLSPAVL 120
Db 57 AACHYKSRVVRMGEHGHMTEGNEQFIDSSRVIRHPQYDSYNDINDIMLIKLSKPAVL 116
QY 121 ARVATISLPRAAAPGTCQLISGWNTLSSGTNYPELLQCLDAPILTAQCEASYPGQI 180
Db 117 QYQTVLALPSSCAPAGTMCVLYSGWNTMSNVSG--DKLQCLQIPILSDRCKNSYFGMIT 174
QY 181 ENMICAGFLEGGKDCQSGSGPVVNCNGELQIVSGYGAQKNGPVYTKVCFNFDWI 240
Db 175 ESMFCAGYLEGGKDCQSGSGPVVNCNGELQIVSGYGAQKNGPVYTKVCFNFDWI 234
QY 241 STIA 244
Db 235 STMA 238

RESULT 8
Q91515
ID Q91515 PRELIMINARY; PRT; 237 AA.
AC Q91515;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Trypsinogen (Fragment).
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang K., Gan L., Lee I., Roach J., Hood L.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
DR EMBL; U25747; AAA75001.1; -.
DR HSSP; P35031; IBIT.
DR MEROPS; S01.125; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolyase; Serine protease.
DR NON_TER 1
FT NON_TER 1
SQ SEQUENCE 237 AA; 25726 MW; 30D2DBAAC39080C2 CRC64;

Query Match 67.8%; Score 899.5; DB 13; Length 237;
Best Local Similarity 68.2%; Pred. No. 2.5e-75;

```

```
Matches 161. Conservative 32; Mismatches 42; Indels 1; Gaps 1;
QY 10 LGAATVPTDDDDKIYGGYTCEENSPYQVSLNAGYHFCGSLISDQWVSAACHYKRSI 69
DB 2 IAAAYAPIDEDDKIYGGYCKKNSVAYQVSLNSYHFCGSLVNNWVNSAACHYKRSV 61
QY 70 QVRGEYNIDVLENEQFINSKAVIRHPNYSWILNDIMLIKSPAVLNARVATISLP 129
DB 62 VVRLGEHNIRANEGTEQFISSRVIRHPNYSYNIDNDIMLIKSLKSPATLNQYVQVVALP 121
QY 130 RACAAPTGQCLISGNTLSSGTNYPELLOCLDAPILTOAQCEASYPGQITENMICAGFL 189
DB 122 SSCAAAGTCKVSGWNTMSS-TADRNKLOCLNIPILSDRDCNSYPGMITDAMFCAGYL 180
QY 190 EGGKDSQGGSGGPVVCNGELQGVSWGYCAOKNKGVTYKVCNFVDMWISTIAA 245
DB 181 EGGKDSQGGSGGPVVCNNELQGVVSWGYGCAERDHPGVYAKVCLFNWLESTMAS 236
--RESULT 9
3266
AC O93266 PRELIMINARY; PRT; 242 AA.
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Trypsinogen 2 precursor (EC 3.4.21.4).
GN TR2.
OS Pseudopleuronectes americanus (Winter flounder) (Pleuronectes americanus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidae; Pleuronectidae; Pseudopleuronectes.
OX NCBI_TaxID=8265;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=INTESTINE, PYLORIC CAECA, AND PANCREAS;
RA Douglas S.E., Gallant J.W.
RT "Isolation of cDNAs for Trypsinogen from the Winter Flounder,
RT Pleuronectes americanus."
RL J. Mar. Biotechnol. 0:0-0(1998).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AF012463; AAC32752.1; -.
DR HSP; P00763; IDPO.
DR MEROPS; S01.125; -.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Signal.
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 21 242 TRYPSINOGEN 2.
SQ SEQUENCE 242 AA; 26180 MW; 08D2A834FB289080 CRC64;
Query Match 67.6%; Score 897; DB 13; Length 242;
Best Local Similarity 64.5%; Pred. No. 4.4e-75;
Matches 158; Conservative 40; Mismatches 43; Indels 4; Gaps 2;
QY 1 MNPLILAFGLAAVATPTDDDDKIYGGYTCEENSPYQVSLNAGYHFCGSLISDQWVVS 60
DB 1 MRSVLVLLIGAFAFAL---EDDKIVGGYECTPHSQAHQVSLNSGYHFCGSLVNNWVVS 57
QY 61 AAHCYKSRIOVRLGEYNIDVLENEQFINSKAVIRHPNYSWILNDIMLIKSPAVLN 120
DB 58 AAHCYKSRVIRGMEGHKIRVNEGTEQFSSSRVIRHPNYSWILNDIMLIKSPATLN 117
QY 121 ARVATISLPACAAPTGQCLISGNTLSSGTNYPELLOCLDAPILTOAQCEASYPGQIT 180
--
Matches 161. Conservative 32; Mismatches 42; Indels 1; Gaps 1;
QY 10 LGAATVPTDDDDKIYGGYTCEENSPYQVSLNAGYHFCGSLISDQWVSAACHYKRSI 69
DB 2 IAAAYAPIDEDDKIYGGYCKKNSVAYQVSLNSYHFCGSLVNNWVNSAACHYKRSV 61
QY 70 QVRGEYNIDVLENEQFINSKAVIRHPNYSWILNDIMLIKSPAVLNARVATISLP 129
DB 62 VVRLGEHNIRANEGTEQFISSRVIRHPNYSYNIDNDIMLIKSLKSPATLNQYVQVVALP 121
QY 130 RACAAPTGQCLISGNTLSSGTNYPELLOCLDAPILTOAQCEASYPGQITENMICAGFL 189
DB 122 SSCAAAGTCKVSGWNTMSS-TADRNKLOCLNIPILSDRDCNSYPGMITDAMFCAGYL 180
QY 190 EGGKDSQGGSGGPVVCNGELQGVSWGYCAOKNKGVTYKVCNFVDMWISTIAA 245
DB 181 EGGKDSQGGSGGPVVCNNELQGVVSWGYGCAERDHPGVYAKVCLFNWLESTMAS 236
--RESULT 9
3266
AC O93266 PRELIMINARY; PRT; 242 AA.
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Trypsinogen 2 precursor (EC 3.4.21.4).
GN TR2.
OS Pseudopleuronectes americanus (Winter flounder) (Pleuronectes americanus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidae; Pleuronectidae; Pseudopleuronectes.
OX NCBI_TaxID=8265;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=INTESTINE, PYLORIC CAECA, AND PANCREAS;
RA Douglas S.E., Gallant J.W.
RT "Isolation of cDNAs for Trypsinogen from the Winter Flounder,
RT Pleuronectes americanus."
RL J. Mar. Biotechnol. 0:0-0(1998).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AF012463; AAC32752.1; -.
DR HSP; P00763; IDPO.
DR MEROPS; S01.125; -.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Signal.
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 21 242 TRYPSINOGEN 2.
SQ SEQUENCE 242 AA; 26180 MW; 08D2A834FB289080 CRC64;
Query Match 67.6%; Score 897; DB 13; Length 242;
Best Local Similarity 64.5%; Pred. No. 4.4e-75;
Matches 158; Conservative 40; Mismatches 43; Indels 4; Gaps 2;
QY 1 MNPLILAFGLAAVATPTDDDDKIYGGYTCEENSPYQVSLNAGYHFCGSLISDQWVVS 60
DB 1 MRSVLVLLIGAFAFAL---EDDKIVGGYECTPHSQAHQVSLNSGYHFCGSLVNNWVVS 57
QY 61 AAHCYKSRIOVRLGEYNIDVLENEQFINSKAVIRHPNYSWILNDIMLIKSPAVLN 120
DB 58 AAHCYKSRVIRGMEGHKIRVNEGTEQFSSSRVIRHPNYSWILNDIMLIKSPATLN 117
QY 121 ARVATISLPACAAPTGQCLISGNTLSSGTNYPELLOCLDAPILTOAQCEASYPGQIT 180
```

```
DB 118 QYVKTVALPSSCAPAGTMTCKVSGWNTMSSADN-GDLQCLDIPILSFSDCNNAYPGMIT 176
QY 181 ENMICAGFLEGGKDSQGGSGGPVVCNGELQGVSWGYCAOKNKGVTYKVCNFVDMWIO 240
DB 177 DSMFCAGYLEGGKDSQGGSGGPVVCNGELQGVVSWGYGCAERDHPGVYAKVCLFNWLE 236
QY 241 STIAA 245
DB 237 STMAS 241
RESULT 10
Q98THO
ID Q98THO PRELIMINARY; PRT; 240 AA.
AC Q98THO;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Trypsinogen.
OS Engraulis japonicus (Japanese anchovy).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Engraulidae;
OC Engraulis.
OX NCBI_TaxID=42892;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PYLORIC CAECA;
RA Watabe S., Ahsan M.N., Funabara D.;
RT "Anchovy trypsinogen mRNA."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB041929; BAB40329.1; -.
DR HSP; P00763; IDPO.
DR MEROPS; S01.125; -.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 240 AA; 26026 MW; 299B119BFF071464 CRC64;
Query Match 66.8%; Score 887; DB 13; Length 240;
Best Local Similarity 63.8%; Pred. No. 3.7e-74;
Matches 157; Conservative 41; Mismatches 42; Indels 6; Gaps 2;
QY 1 MNPLILAFGLAAVATPTDDDDKIYGGYTCEENSPYQVSLNAGYHFCGSLISDQWVVS 60
DB 1 MRSVLVLLIGAFAFAL---EDDKIVGGYECTPHSQAHQVSLNSGYHFCGSLVNNWVVS 56
QY 61 AAHCYKSRIOVRLGEYNIDVLENEQFINSKAVIRHPNYSWILNDIMLIKSPAVLN 120
DB 57 AAHCYKSRVIRGMEGHKIRVNEGTEQFIDSSRVIRHPNYSYNIDNDIMLIKSPATLN 116
QY 121 ARVATISLPACAAPTGQCLISGNTLSSGTNYPELLOCLDAPILTOAQCEASYPGQIT 180
DB 117 QYVQVVALPSSCAPAGTMTCKVSGWNTMSSVSG-DKLOCLQIPILSDRDCNSYPGMIT 174
QY 181 ENMICAGFLEGGKDSQGGSGGPVVCNGELQGVSWGYCAOKNKGVTYKVCNFVDMWIO 240
DB 175 DAMFCAGYLEGGKDSQGGSGGPVVCNGELQGVVSWGYGCAERDHPGVYAKVCLFNWLE 234
QY 241 STIAA 246
DB 235 STMAS 240
RESULT 11
```

```

Q9W7Q7
ID Q9W7Q7 PRELIMINARY; PRT; 242 AA.
AC Q9W7Q7;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Trypsinogen 1.
OS Paralichthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidei; Paralichthyidae; Paralichthys.
OX NCBI_TaxID=8255;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PANCREAS;
RA Suzuki T., Srivastava A.S., Kurokawa T.;
RT "Japanese flounder mRNA for trypsinogen 1.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
-2- EMBL; AB029750; BAA82362.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.125; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 242 AA; 26548 MW; 6DA722C80BC194A2 CRC64;

Query Match 66.2%; Score 879; DB 13; Length 242;
Best Local Similarity 63.3%; Pred No. 2e-73;
Matches 155; Conservative 41; Mismatches 45; Indels 4; Gaps 2;

QY 1 MNPLILAFGLGAATPTDDDDKIVGGYTCEENSVYQVSLNAGYHFCGSLSDQWVVS 60
DB 1 MRSLVFLLIGLGAFAFAM---EDDKIVGGYECTPYSPHQVSLNSGYHFCGSLVNNWVS 57
QY 61 AAHCYKSRIOVRLGEYNIDVLENGEOPFINSKIVIRHPNYSWILDNDIMLIKSSPAVLN 120
DB 58 AAHCYKSRVEVRGEHGIKINEGTEQFISSEIRVIRHPNYSYINNDIMLIKREPAVLN 117
-2- 121 ARVATISLPACAAPTQCLISGNTLSGGTNYPELLOCLDAPILTQAOCEASYPGQIT 180
-3- 118 QYQVPALPTSCAPAGTMCVTSGWNTMSSTAN-RDMLOCLDLPILSDRCENSYPGMIT 176
QY 181 ENMICAGFLEGGKDCQSGGPGVPCNGELQGVSWGYGCAQKNKPGYTKVCNFVDWIQ 240
DB 177 PAMFCAGYLEGGKDCQSGGPGVPCNGELQGVSWGYGCAERDHPGVYARVCIFIDWLE 236
QY 241 STTAA 245
DB 237 TTMAAS 241

RESULT 12
Q8QGW3
ID Q8QGW3 PRELIMINARY; PRT; 244 AA.
AC Q8QGW3;
DT 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Trypsinogen (EC 3.4.21.4).
OS Anguilla japonica (Japanese eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
OC Anguilla.

```

```

OX NCBI_TaxID=7937;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PANCREAS;
RA Kurokawa T., Suzuki T., Ohta H., Kagawa H., Tanaka H., Unuma T.;
RT "Expression of pancreatic enzyme genes during the early larval stage of Japanese eel, Anguilla japonica.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB070720; BAB85634.1; -.
KW Hydrolase.
SQ SEQUENCE 244 AA; 26317 MW; 0EB3B68E8706D52D CRC64;

Query Match 65.7%; Score 872; DB 13; Length 244;
Best Local Similarity 63.6%; Pred No. 9.1e-73;
Matches 157; Conservative 38; Mismatches 46; Indels 6; Gaps 3;

QY 1 MNPLILAFGLGAATPTDDDDKIVGGYTCEENSVYQVSLNAGYHFCGSLSDQWVVS 60
DB 1 MRSLVFLLIGLGAVAL---DDDKIVGGYCEPHSQVQASLNAGYHFCGSLVNNWVS 57
QY 61 AAHCYK--SRIOVRLGEYNIDVLENGEOPFINSKIVIRHPNYSWILDNDIMLIKSSPAV 118
DB 58 AAHCYKSPSRLEVRLEHIGLNEGTEQFICASHVIRPNPNYSWLDSDIMLIKSKPAT 117
QY 119 LNARVATISLPACAAPTQCLISGNTLSGGTNYPELLOCLDAPILTQAOCEASYPGQ 178
DB 118 LNSYQVPALPTSCAPAGTMCRTGWTMPNPAVS-GDKLOCLEIPILSEDCNSYPGM 176
QY 179 ITENMICAGFLEGGKDCQSGGPGVPCNGELQGVSWGYGCAQKNKPGYTKVCNFVDW 238
DB 177 ITSTMFCAGYLEGGKDCQSGGPGVPCNGELQGVSWGYGCAEQNHQPGVYPKVCMFSDW 236
QY 239 IOSTTAA 245
DB 237 LRTTMAAS 243

RESULT 13
Q9W7Q6
ID Q9W7Q6 PRELIMINARY; PRT; 238 AA.
AC Q9W7Q6;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Trypsinogen 2 (Fragment).
OS Paralichthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidei; Paralichthyidae; Paralichthys.
OX NCBI_TaxID=8255;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PANCREAS;
RA Suzuki T., Srivastava A.S., Kurokawa T.;
RT "Japanese flounder mRNA for trypsinogen 2.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
DR EMBL; AB029751; BAA82363.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.125; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolase; Serine protease.
FT NON_TER 1
SQ SEQUENCE 238 AA; 26071 MW; F2B8908085B8D062 CRC64;

```

```

Query Match          65.7%; Score 871.5; DB 13; Length 238;
Best Local Similarity 63.6%; Pred. No. 9.9e-73;
Matches 154; Conservative 41; Mismatches 42; Indels 5; Gaps 2;

QY 4 LLLIALLFGLAAVATPTDDDDKIVGGYTCSENSVPYQVSLNAGYHFCGSLSDQWVSAH 63
DB 1 LVPILLIGAAFAF---EDDKIVGGYECTPYSPHQVSLNSGYHFCGSLVNNWVSAH 57

QY 64 CYKSRIOVRLGEYNIDVLENEQFINSKAVIRHPNYSWILNDIMLIKSSPAVLNARV 123
DB 58 CYKSRVEVRIGENLRVETEYFISSSRVIRHPNYSYNINNDIMLIKSEPAFLNQY 117

QY 124 ATISLPACAAAPGTQCLISGWGNTLSSGTNYPELLOCLDAPILTQAOCEASYPGOITENM 183
DB 118 QPVALPTSCAPACTMCTVSGWGTMSSTD--SRLQCLDLPILSERDCENSYPGMITNAM 175

QY 184 ICAGFLEGGKDCSGDGGPVVNGELQIVSWGYGCAQKNKPGYVTKVCFVDMQSTI 243
DB 176 FCAGYLEGGKDCSGDGGPVVNGELQIVSWGYGCAQKNKPGYVTKVCFVDMQSTI 235

-/- 244 AA 245
DB 236 SS 237

RESULT 14
O92099 PRELIMINARY; PRT; 242 AA.
AC O92099;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Trypsin precursor (EC 3.4.21.4).
OS Paranthothenia magellanica (Maori cod).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes;
OC Notothenioidae; Nototheniidae; Paranthothenia.
OX NCBI_TaxID=37005;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PYLORIC CAECA;
RX MEDLINE=97104330; PubMed=8948488;
RA Genicot S., Rentier-Delrue F., Edwards D., Van Beemmen J., Gerday C.;
RT "Trypsin and trypsinogen from an antarctic fish: molecular basis of
RT cold adaptation.";
RL Biochim. Biophys. Acta 1298:45-57(1996).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; X82223; CAA57701.1; -
DR HSP; P00763; IDPO.
DR MEROPS; S01.125; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Signal.
FT SIGNAL 1 13 POTENTIAL.
FT CHAIN 16 247 TRYPSIN A3.
SQ SEQUENCE 242 AA; 26201 MW; 3F4DE7CE80C4477C CRC64;

Query Match          64.4%; Score 854; DB 13; Length 242;
Best Local Similarity 62.3%; Pred. No. 4.2e-71;
Matches 152; Conservative 42; Mismatches 46; Indels 4; Gaps 2;

QY 1 MNPILLIALLFGLAAVATPTDDDDKIVGGYTCSENSVPYQVSLNAGYHFCGSLSDQWVVS 60
DB 1 MRSLVFVLLIGAAFAF---EEDKIVGKGCSPYSPHQVSLNSGYHFCGSLVNNWVVS 57

Query Match          63.2%; Score 838.5; DB 13; Length 247;
Best Local Similarity 63.3%; Pred. No. 1.2e-69;
Matches 157; Conservative 25; Mismatches 63; Indels 3; Gaps 2;

QY 1 MNPILLIALLFGLAAVATPTDDDDKIVGGYTCSENSVPYQVSLNAGYHFCGSLSDQWVVS 60
DB 1 MHGLIALLVGVAAAPYMYEDHIVGSECAAHSPQWQVSLNIGYHFCGSLNSQWVVS 60

QY 61 AAHCYK--SRIOVRLGEYNIDVLENEQFINSKAVIRHPNYSWILNDIMLIKSSPAV 118
DB 61 AAHCYOTASIRISVIGENIRFVNEGTQQIOAKAIOHPQYNSTWIDIMLIKSSPAT 120

QY 119 LNARVATISLPACAAAPGTQCLISGWGNTLSSGTNYPELLOCLDAPILTQAOCEASYPGO 178
DB 121 LNQYQAIALPSSCVNTGVMTISGWGETQTS--VGSFVLMCVQAPVLSDTSCRNSTPGD 179

```

Qy	179	ITNNMICAGLEGGKDCQDSDGGPVCNGELQGIIVSWGCAQKPKGVYTKVCNFVDW 438
Db	180	ITNNMICAGLEGGKDCQDSDGGPVCNGELQGIIVSWGCAQKPKGVYTKVCNFVDW 438
Qy	239	IQSTIAAN 246
Db	240	IAQTIAAN 247

Search completed: June 20, 2003, 15:21:52
Job time : 85 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 15:06:31 ; Search time 80 seconds
(without alignments)
411.411 Million cell updates/sec

Title: US-09-762-277A-1

Perfect score: 1327

Sequence: 1 MNPLLILAFGLAAVATPTDD.....VYTKVCFVDFWQSTIRANS 247

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1327	100.0	247	21	AAW78974
2	1111	83.7	247	18	AAW08475
3	1101	83.0	246	19	AAW64260
4	1097	82.7	247	23	AAU87693
5	1085	81.6	247	21	AAU81321
6	1063.5	80.1	247	21	AAU35701
7	1058	79.7	241	21	AAU21316
8	1058	79.7	247	9	AAU81243
9	1058	79.7	247	16	AAU82703
10	1036	78.1	240	19	AAW57740
					Canine cationic tr
					Porcine trypsinog
					Human amyloid beta
					Human pancreatic t
					Human trypsinogen
					Human trypsin hL a
					Human spleen tryps
					Human pancreatic t
					Trypsinogen-like p

11	1026	77.3	230	20	AAW93488
12	1026	77.3	246	21	AAW78975
13	1002	75.5	225	20	AAW89503
14	997	75.1	224	20	AAW31160
15	978	73.7	229	21	AAW77494
16	978	73.7	231	21	AAW91926
17	978	73.7	231	22	AAW80953
18	974	73.4	233	21	AAW91925
19	967	72.9	230	15	AAW53638
20	949	71.5	223	20	AAW81767
21	949	71.5	223	21	AAW69973
22	949	71.5	224	15	AAW53637
23	947	71.4	223	23	ABB83322
24	907	68.3	241	23	ABB04645
25	887	66.8	240	23	ABB04644
26	734	55.3	238	20	AAW08894
27	727	54.8	225	22	AAW31579
28	697.5	52.6	254	23	AAU74758
29	650	49.0	151	22	AAW68927
30	628.5	47.4	289	21	AAW36483
31	628.5	47.4	289	22	AAW67543
32	614	46.3	251	21	AAW21306
33	605	45.6	237	21	AAW21305
34	602.5	45.4	288	21	AAW36482
35	602.5	45.4	288	22	AAW67542
36	578	43.6	293	20	AAW30524
37	578	43.6	293	20	AAW38412
38	578	43.6	293	20	AAW38426
39	578	43.6	293	20	AAW16777
40	578	43.6	293	21	AAW21296
41	578	43.6	293	21	AAW66726
42	578	43.6	293	22	AAU12399
43	578	43.6	293	22	AAW65249
44	578	43.6	293	23	AAU81966
45	578	43.6	322	20	AAW14072

ALIGNMENTS

RESULT 1

AAW78974
ID AAW78974 standard; protein; 247 AA.
AC AAW78974;
DT 05-JUN-2000 (first entry)
DE Canine cationic trypsinogen amino acid sequence.
KW Cationic trypsinogen; dog; monoclonal antibody production; detect;
KW diagnose; pancreatitis; pancreatic cancer; renal insufficiency;
KW extrapancreatic hyposecretion.
OS Canis familiaris.
PN WO200009739-A1.
XX WO200009739-A1.
PD 24-FEB-2000.
XX 09-AUG-1999; 99WO-JP04299.
XX 10-AUG-1999; 98JP-0236609.
XX 10-MAR-1998; 95JP-0063990.
XX (FUJY) FUJI YAKUHIIN KOGYO KK.
XX Waritani T, Ashida Y, Yamada T;
XX WPI; 2000-206018/18.
XX Anti-canine trypsin monoclonal antibody, useful as reagent for quick
XX and accurate detection and quantitation of trypsin and/or trypsin-like

Human TRYI trypsin
Canine anionic try
Human trypsin seri
Human trypsin seri
Bovine trypsinogen
Recombinant trypsi
Bovine met-phe-try
Trypsinogen analog
Bovine trypsinogen
Bovine TRYP peptid
TRYP protein. Bos
Bovine trypsin. B
Partial trypsin se
Engraulis japonicu
Engraulis japonicu
Chimeric serine pr
Amino acid sequenc
Human protease PRT
Human protease dom
Fusion gene with h
Amino acid sequenc
Human KLK-16 prote
Human KLK-16 prote
Fusion gene with h
Amino acid sequenc
Human PDSP-1 prote
Human secreted pro
Human secreted pro
Human keratinocyte
Human KLK-L2 prote
Membrane-bound pro
Human PRO1132 poly
Human PRO1132 (UNQ
Human PRO1132. HO
Human BS247 specif

PT immunoreactants in various forms in diagnosis e.g. of pancreatic
PT diseases -

Claim 3; Page 63-64; 67pp; Japanese.

XX This sequence represents the canine cationic trypsinogen amino acid
CC sequence. The protein was isolated from the canine pancreas. The
CC invention relates to monoclonal antibodies with specificity against
CC canine trypsin, or canine trypsin-related substances. The antibodies are
CC highly specific and can be used as reagent for quick and accurate
CC detection and quantitation of canine trypsin and canine trypsin-like
CC immunoreactants in various forms. The antibodies can be used in the
CC diagnosis of diseases such as pancreatitis, pancreatic cancer, renal
CC insufficiency and extrapancreatic hyposecretion.

Query Match 100.0%; Score 1327; DB 21; Length 247;
Best Local Similarity 100.0%; Pred. No. 6.8e-114;
watches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

XX PD 14-SEP-2000.
 XX PF 09-MAR-2000; 2000WO-CA00258.
 XX PR 11-MAR-1999; 99US-0124260.
 XX PR 01-APR-1999; 99US-0127386.
 XX PR 21-JUL-1999; 99US-0144919.
 XX PA (MOUN) MOUNT SINAI HOSPITAL.
 XX PI Yousef GM, Diamandis EP;
 XX WPI; 2000-587440/55.
 XX DR New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L
 XX PT protein mediated disorders, especially cancer.
 XX PS Example 4; Fig 17; 184pp; English.

The present sequence is human trypsinogen, a member of the serine
 protease family. Kallikreins and kallikrein-like proteins are a
 subgroup of the serine protease enzyme family. They catalyze the
 selective cleavage of specific polypeptide precursors to release
 peptides with potent biological activity. Nucleic acids encoding
 kallikrein-like proteins KLK-L1, KLK-L2, KLK-L3, KLK-L4, KLK-L5 and
 KLK-L6 have been isolated. The proteins are useful in the treatment,
 monitoring and diagnosis of cancers, especially prostate cancer. They
 can also be used to identify a substance that can associate with or
 mediate the biological activity of the proteins. Antibodies can
 be used to treat conditions mediated by the kallikrein-like proteins.

XX SQ Sequence 247 AA;

Query Match 81.8%; Score 1085; DB 21; Length 247;
 Best Local Similarity 80.2%; Pred. No. 1.2e-91;
 Matches 198; Conservative 25; Mismatches 24; Indels 0; Gaps 0;

OY 1 MNPLLIALLGAATPTDDDDKIVGGYTCENSVYQVSLNAGYHFCGSLISDQWVS 60
 DB 1 MNPLLIITFVNAALAAFFDDDDKIVGGYTCENSVYQVSLNAGYHFCGSLISDQWVS 60
 OY 61 AAHCYKSRIOVRLGEYINIDVLEGNQFINSARVIRHPNYSWILNDIMLIKLSPPAVLN 120
 DB 61 AGHCYKSRIOVRLGEYINIDVLEGNQFINSARVIRHPNYSWILNDIMLIKLSPPAVLN 120
 OY 121 ARVATISLPACAPGTQCLISGNGNTLSSTNTPELLOCLDAPILTOAQCEASYPGQIT 180
 DB 121 ARVETISLPTAPPATGTKCLISGNGNTASSGADYDDELQCLDAPVLSQAKCEASYPGKIT 180
 OY 181 ENMICAGFLEGGKDCQDGGPVVVCNGELGIVSWGACQAKNKPQVYTKVCFNVDWIQ 240
 DB 181 SNMFCVFLGEGKDCQDGGPVVVCNGELGIVSWGACQAKNKPQVYTKVCFNVDWIQ 240
 OY 241 STIAANS 247
 DB 241 NTIAANS 247

RESULT 6
 AAB35701
 ID AAB35701 standard; Protein; 247 AA.
 AC AAB35701;
 XX 14-FEB-2001 (first entry)
 DT Human trypsin hL amino acid sequence.
 DE Human; trypsin hL; serine protease; lung disease model animal.
 KW Homo sapiens.
 OS
 XX

PN JP2000253887-A.
 XX 19-SEP-2000.
 XX PF 11-MAR-1999; 99JP-0065337.
 XX PR 11-MAR-1999; 99JP-0065337.
 XX PA (TTPH-) TT PHARMA KK.
 XX WPI; 2000-658814/64.
 XX DR N-PSDB; AAC66182.
 XX PT Novel gene encoding a serine protease and its protein used to screen
 XX PT for serine protease inhibitors and to prepare lung disease animal
 XX PS models
 PS Claim 2; Page 3-4; 17pp; Japanese.

Nucleotide sequence AAC66182 encodes human trypsin hL AAB35701, a serine
 protease. The invention relates to the human hL gene and protein
 sequences, to a recombinant vector containing the nucleotide sequence,
 and a host cell containing the vector. Human trypsin hL can be used for
 screening for serine protease inhibitors, in the preparation of a lung
 disease model animal, and for the development of an index marker of lung
 diseases caused by an anti-trypsin hL antibody.

XX SQ Sequence 247 AA;

Query Match 80.1%; Score 1063.5; DB 21; Length 247;
 Best Local Similarity 76.1%; Pred. No. 1.1e-89;
 Matches 188; Conservative 31; Mismatches 27; Indels 1; Gaps 1;

OY 1 MNPLLIALLGAATPTDDDDKIVGGYTCENSVYQVSLNAGYHFCGSLISDQWVS 59
 DB 1 MKTLIFLAFGLAAVALPLDDDDKIVGGYTCORNALPYQVSLNAGYHFCGSLINQWVS 60
 OY 60 SAHCYKSRIOVRLGEYINIDVLEGNQFINSARVIRHPNYSWILNDIMLIKLSPPAVLN 119
 DB 61 SAHCYKSRIOVRLGEYINIDVLEGNQFINSARVIRHPNYSWILNDIMLIKLSPPAVLN 120
 OY 120 NARVATISLPACAPGTQCLISGNGNTLSSTNTPELLOCLDAPILTOAQCEASYPGQI 179
 DB 121 NSRVETVALPRSCPSAGTRCLVSGWNTLSSTNTPELLOCLDAPVLSDSSTSSYPGKI 180
 OY 180 TENMICAGFLEGGKDCQDGGPVVVCNGELGIVSWGACQAKNKPQVYTKVCFNVDWI 239
 DB 181 TSNMFCVFLGEGKDCQDGGPVVVCNGELGIVSWGACQAKNKPQVYTKVCFNVDWI 240
 OY 240 QSTIAAN 246
 DB 241 QOTIAAN 247

RESULT 7
 AAB21316
 ID AAB21316 standard; Protein; 241 AA.
 AC AAB21316;
 XX 02-FEB-2001 (first entry)
 DT Human trypsinogen.
 DE Human; trypsinogen.
 KW Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6; trypsinogen;
 KW kallikrein-like protein; serine protease;
 KW cytostatic; cancer; prostate cancer.
 OS Homo sapiens.
 XX WO200053776-A2.
 XX 14-SEP-2000.
 PD

XX 09-MAR-2000; 2000WO-CA00258.
 XX 11-MAR-1999; 99US-0124260.
 PR 01-APR-1999; 99US-0127386.
 PR 21-JUL-1999; 99US-0144919.
 XX (MOUN) MOUNT SINAI HOSPITAL.
 XX Yousef GM, Diamandis EP;
 PI WPI; 2000-587440/55.
 DR New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L
 PT protein mediated disorders, especially cancer. -
 PT Example 3; Fig 9; 184pp; English.
 PS The present sequence is human trypsinogen, a member of the kallikrein
 XX multi-gene family. Kallikreins and kallikrein-like proteins are a
 CC subgroup of the serine protease enzyme family. They catalyze the
 CC selective cleavage of specific polypeptide precursors to release
 CC peptides with potent biological activity. Nucleic acids encoding
 CC kallikrein-like proteins KLK-L1, KLK-L2, KLK-L3, KLK-L4, KLK-L5 and
 CC KLK-L6 have been isolated. The proteins are useful in the treatment,
 CC monitoring and diagnosis of cancers, especially prostate cancer. They
 CC can also be used to identify a substance that can associate with or
 CC mediate the biological activity of the proteins. Antibodies can
 CC be used to treat conditions mediated by the kallikrein-like proteins.
 XX Sequence 241 AA;
 SQ Query Match 79.7%; Score 1058; DB 21; Length 241;
 Best Local Similarity 79.7%; Pred. No. 3.5e-89;
 Matches 192; Conservative 25; Mismatches 24; Indels 0; Gaps 0;
 QY 1 MNPLILFLAFLGAATPTDDDDKIVGGYTCENSPYQVSLNAGYHFCGSLISDQWVVS 60
 DB 1 MNPLILFLAFLGAATPTDDDDKIVGGYTCENSPYQVSLNAGYHFCGSLISDQWVVS 60
 QY 61 AAHCYKSRIOVRLGEYNDVLEGNQFINSKAVIRHPNYSWILDNDIMLIKLSRAVIN 120
 DB 61 AAHCYKSRIOVRLGEYNDVLEGNQFINSKAVIRHPNYSWILDNDIMLIKLSRAVIN 120
 QY 121 ARVATISLPRACAPGTQCLISGWNTLSSTGTYNYPQLQCLDAPILTOACEASYPGQIT 180
 DB 121 ARVSTISLPTAPPATGTCTKLSGWNTASSGADYDEQLQCLDAPVLSQAKCEASYPGKIT 180
 QY 181 ENMICAGFLEGKDCSQGDSGGPVVNCNGELQGVSWGCGCAQKNKPGVYTKVNFVDWIK 240
 DB 181 SNMFCVGFLEGKDCSQGDSGGPVVNCNGELQGVSWGCGCAQKNKPGVYTKVNFVDWIK 240
 QY 241 S 241
 DB 241 N 241
 RESULT 8
 AAP81243
 ID AAP81243 standard; protein; 247 AA.
 AC AAP81243;
 XX 07-NOV-1990 (first entry)
 DT Human spleen trypsin III (trypsinogen III).
 DE Human spleen plasminogen; trauma lesions.
 KW Homo sapiens.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FH Misc-difference 1..22
 FT

FT /label=deletion
 FT /note="optional"
 FT Misc-difference 1..23
 FT /label=deletion
 FT /note="optional"
 XX JP63160582-A.
 PN 04-JUL-1988.
 XX 25-DEC-1986; 86JP-0307770.
 XX 25-DEC-1986; 86JP-0307770.
 PR (SANY) SANKYO KK.
 XX WPI; 1988-224890/32.
 DR N-PSDB; AAN81633.
 DR Human spleen trypsin - used to treat lesions of trauma, without
 XX hypersensitive allergic side effects.
 PT Claim 2+1; Page 1; 9pp; Japanese.
 PS Expression vectors E.coli LE 392 and YA 21 are preferable for mass
 XX production, and animal cells or B.subtilis are suitable for the
 CC production of an enzyme of similar activity to that of natural human
 CC spleen trypsinogen. Culturing the recombinant cells produced the
 CC desired trypsin as insoluble protein in inclusion bodies and the
 CC trypsin was then isolated and purified.
 CC The product is used in the treatment of lesions or trauma, e.g.
 CC burns, gangrene, abscesses, injury etc.
 XX SQ Sequence 247 AA;
 Query Match 79.7%; Score 1058; DB 9; Length 247;
 Best Local Similarity 78.5%; Pred. No. 3.6e-89;
 Matches 194; Conservative 27; Mismatches 26; Indels 0; Gaps 0;
 QY 1 MNPLILFLAFLGAATPTDDDDKIVGGYTCENSPYQVSLNAGYHFCGSLISDQWVVS 60
 DB 1 MNPLILFLAFLGAATPTDDDDKIVGGYTCENSPYQVSLNAGYHFCGSLISDQWVVS 60
 QY 61 AAHCYKSRIOVRLGEYNDVLEGNQFINSKAVIRHPNYSWILDNDIMLIKLSRAVIN 120
 DB 61 AAHCYKSRIOVRLGEYNDVLEGNQFINSKAVIRHPNYSWILDNDIMLIKLSRAVIN 120
 QY 121 ARVATISLPRACAPGTQCLISGWNTLSSTGTYNYPQLQCLDAPILTOACEASYPGQIT 180
 DB 121 ARVSTISLPTAPPATGTCTKLSGWNTASSGADYDEQLQCLDAPVLSQAKCEASYPGKIT 180
 QY 181 ENMICAGFLEGKDCSQGDSGGPVVNCNGELQGVSWGCGCAQKNKPGVYTKVNFVDWIK 240
 DB 181 SNMFCVGFLEGKDCSQGDSGGPVVNCNGELQGVSWGCGCAQKNKPGVYTKVNFVDWIK 240
 QY 241 STIAANS 247
 DB 241 DTIAANS 247
 RESULT 9
 AAR82703
 ID AAR82703 standard; Protein; 247 AA.
 XX AC AAR82703;
 XX 19-MAR-1996 (first entry)
 DT Human pancreatic trypsin III.
 XX Pancreatic trypsin III; trypsinogen; human; hydrolysis; ds.
 KW Pancreatic trypsin III; trypsinogen; human; hydrolysis; ds.
 XX

OS Homo sapiens.
 XX Key Location/Qualifiers
 FH Peptide 1..45
 FT /label= signal_peptide
 FT
 XX JP07184655-A.
 XX PN
 XX PD
 XX PF 25-JUL-1995.
 XX PF 25-DEC-1986; 86JP-0307770.
 XX PR 25-DEC-1986; 86JP-0307770.
 XX PR 25-DEC-1986; 86JP-0311512.
 XX PA (SANY) SANKYO CO LTD.
 XX WPI; 1995-287966/38.
 XX DR N-PSDB; AAT03999, AAT04000, AAT04001.
 Novel human pancreatic trypsin III - can be easily produced by recombinant methods
 XX Claim 1; Page 10; 11pp; Japanese.
 XX CC AAT03999-T04001 are all human cDNA sequences which code for pancreatic trypsin III (AAR87203), the sequences differ only in their stop codons. The cDNA molecules can be used in the recombinant production of trypsin which can be used as a drug to treat diseases wherein trypsin production is impaired.
 XX CC trypsin production is impaired.
 XX SQ Sequence 247 AA;
 Query Match 79.7%; Score 1058; DB 16; Length 247;
 Best Local Similarity 78.5%; Pred. No. 3.6e-89;
 Matches 194; Conservative 27; Mismatches 26; Indels 0; Gaps 0;
 QY 1 MNPLLLAFGLAAVATPTDDDDKIVGGYTCEENSVYQVSLNAGYHFCGGLISDQWVVS 60
 DB 1 MNPFLLLAFGLAAVATPTDDDDKIVGGYTCEENSVYQVSLNAGYHFCGGLISDQWVVS 60
 QY 61 AAHCYKSRIOVRGEYNIDVLENEQFINSAKVIHNPNTNSWILDNDIMLIKLSPPAVLN 120
 DB 61 AAHCYKTRIOVRGEHNIKVLNEQFINAAKIIIRHPKYNRDTLDNDIMLIKLSPPAVLN 120
 QY 121 ARVATISLPACAPGTCQLISGNTLSGTYNPELLOCLDAPILTOAQCEASYPGQIT 180
 DB 121 ARVSTISLTPAPPAAGTECLISGNTLSFGADYPDELKCLDAPVLRAEACRACPGKIT 180
 QY 181 ENMICAGFLEGGKDCQSGGPPVVCNGELQGVSWGCAOKNPGVYTKVCFNFDWIO 240
 DB 181 NSMFCVGFLEGGKDSWKRSGGPPVVCNGELQGVSWGHCANRNPVYTKVYNDWIK 240
 QY 241 STIAANS 247
 DB 241 DTIAANS 247
 RESULT 10
 AAW57740
 ID AAW57740 standard; Protein; 240 AA.
 XX
 AC AAW57740;
 XX
 DT 16-SEP-1998 (first entry)
 DE Trypsinogen-like protein.
 XX Trypsinogen-like protein; human.
 KW Homo sapiens.
 OS
 XX JP10099080-A.

XX 21-APR-1998.
 XX 26-SEP-1996; 96JP-0273923.
 XX 26-SEP-1996; 96JP-0273923.
 XX (SHIS) SHISEIDO CO LTD.
 XX WPI; 1998-289873/26.
 XX DR N-PSDB; AAV24548.
 XX DNA coding for trypsinogen-like protein
 XX Claim 3; Page 4-5; 7pp; Japanese.
 XX This sequence represents a human trypsinogen-like protein, and is encoded by the DNA of the invention.
 XX SQ Sequence 240 AA;
 Query Match 78.1%; Score 1036; DB 19; Length 240;
 Best Local Similarity 80.3%; Pred. No. 3.7e-87;
 Matches 188; Conservative 24; Mismatches 22; Indels 0; Gaps 0;
 QY 14 VATPTDDDDKIVGGYTCEENSVYQVSLNAGYHFCGGLISDQWVVSAAHCYKSRIOVRL 73
 DB 7 VAVPFDDDDKIVGGYTCEENSVYQVSLNAGYHFCGGLISDQWVVSAAHCYKTRIOVRL 66
 QY 74 GEYNIDVLENEQFINSAKVIHNPNTNSWILDNDIMLIKLSPPAVLNARVATISLPAC 133
 DB 67 GEHNIKVLNEQFINAAKIIIRHPKYNRDTLDNDIMLIKLSPPAVLNARVSTISLPTAPP 126
 QY 134 APGTCQLISGNTLSGTYNPELLOCLDAPILTOAQCEASYPGQITENMICAGFLEGGK 193
 DB 127 AAGTECLISGNTLSFGADYPDELKCLDAPVLRAEACRACPGKITNSMFCVGFLEGGK 186
 QY 194 DSCOGDSGPPVVCNGELQGVSWGCAOKNPGVYTKVCFNFDWIOSTIAANS 247
 DB 187 DSCORDSGPPVVCNGELQGVSWGHCANRNPVYTKVYNDWIKDTIAANS 240
 RESULT 11
 AAW93488
 ID AAW93488 standard; Protein; 230 AA.
 XX
 AC AAW93488;
 XX
 DT 11-JUN-1999 (first entry)
 DE Human TRYI trypsinogen variant protein.
 XX TRYI; trypsinogen; autocatalysis; cleavage; zymogenic precursor; protease; peptide hormone; therapeutic protein; treatment; coagulation.
 KW Homo sapiens.
 OS
 XX WO9910503-A1.
 XX 04-MAR-1999.
 XX 12-AUG-1998; 98WO-EP05094.
 XX 15-OCT-1997; 97EP-0117816.
 XX 22-AUG-1997; 97EP-0114513.
 XX (HOFF) ROCHE DIAGNOSTICS GMBH.
 XX Bode W, Hopfner K, Huber R, Kopetzki E;
 WPI; 1999-204669/17.
 XX N-PSDB; AAX23298.

PT Autocatalytically cleavable zymogenic protease precursors - useful
 XX for cleaving fusion proteins and for therapeutic uses
 PS Disclosure; Fig 1; 45pp; German.
 CC This invention describes a method where autocatalytically cleavable,
 CC zymogenic precursors of a protease (in this invention trypsinogen),
 CC have their natural cleavage site replaced by a non-natural,
 CC autocatalytically cleavable site. Such proteases are reagents for
 CC producing peptide hormones and other therapeutic proteins by cleavage
 CC of their fusion proteins, and some also have therapeutic activity
 CC themselves, e.g. thrombin for treatment of coagulation disorders. The
 CC proteases are produced simply and in high yield without the need to
 CC add another protease for cleavage, reducing costs and avoiding
 CC contamination of the final enzyme.
 XX
 SQ Sequence 230 AA;
 Query Match 77.3%; Score 1026; DB 20; Length 230;
 Best Local Similarity 81.2%; Pred. No. 2.9e-86;
 Matches 186; Conservative 23; Mismatches 20; Indels 0; Gaps 0;
 QY 19 DDDKIVGGYCEENSVYQVSLNAGYHFCGSLSDQWVSAAHYKSRIOVRLGEYNI 78
 DB 2 DDDKIVGGYCEENSVYQVSLNAGYHFCGSLSDQWVSAAHYKSRIOVRLGEYNI 61
 QY 79 DYLEGNEOFINSKVIHNPNTNSWILDNDIMLIKLSPPAVLNARVATISLPACAPGQ 138
 DB 62 EVLEGNEOFINAAKTIIRHPQYDRKTLNNDIMLIKLSRAVINARVSTISLPATGPK 121
 QY 139 CLISGWNTLSGTYNPELLOCLDAPILTOACEASYPGOITENNICAGFLEGGKDCOG 198
 DB 122 CLISGWNTASGADYPELOCLDAPILVSOAKCEASYPGKIITSNMFCVLEGGKDCOG 181
 QY 199 DSGPVCVNGELQIVSGYCAQKPKGVYTKVCFNFDWIOSTIAANS 247
 DB 182 DSGPVCVNGLOQVSVSGDGAQKPKGVYTKVYKWKNTIAANS 230
 RESULT 12
 AAY78975
 ID AAY78975 standard; protein; 246 AA.
 AC AAY78975;
 XX
 XX 05-JUN-2000 (first entry)
 DT
 DE Canine anionic trypsinogen amino acid sequence.
 KW Anionic trypsinogen; dog; monoclonal antibody production; detect;
 KW diagnose; pancreatitis; pancreatic cancer; renal insufficiency;
 XX extrapancreatic hyposecretion.
 XX
 OS Canis familiaris.
 XX
 XX WO200009739-A1.
 XX
 XX 24-FEB-2000.
 XX
 XX 09-AUG-1999; 99WO-JP04299.
 XX
 XX 10-AUG-1998; 98JP-0236609.
 XX
 XX 10-MAR-1999; 99JP-0063990.
 XX
 XX (FUJY) FUJI YAKUHI KOGYO KK.
 XX
 XX Waritani T, Ashida Y, Yamada T;
 XX
 XX WPI; 2000-206018/18.
 XX
 XX Anti-canine trypsin monoclonal antibody, useful as reagent for quick
 XX and accurate detection and quantitation of trypsin and/or trypsin-like
 XX immunoreactants in various forms in diagnosis e.g. of pancreatic

PT diseases
 XX
 PS Claim 3; Page 64-65; 67pp; Japanese.
 CC
 CC This sequence represents the canine anionic trypsinogen amino acid
 CC sequence. The protein was isolated from the canine pancreas. The
 CC invention relates to monoclonal antibodies with specificity against
 CC canine trypsin, or canine trypsin-related substances. The antibodies are
 CC highly specific and can be used as reagent for quick and accurate
 CC detection and quantitation of canine trypsin and canine trypsin-like
 CC immunoreactants in various forms. The antibodies can be used in the
 CC diagnosis of diseases such as pancreatitis, pancreatic cancer, renal
 CC insufficiency and extrapancreatic hyposecretion.
 XX
 SQ Sequence 246 AA;
 Query Match 77.3%; Score 1026; DB 21; Length 246;
 Best Local Similarity 75.2%; Pred. No. 3.1e-86;
 Matches 185; Conservative 23; Mismatches 38; Indels 0; Gaps 0;
 QY 1 MNPILLAFGLAAVATPTDDDKIVGGYTCCEENSVYQVSLNAGYHFCGSLSDQWVVS 60
 DB 1 MKTFIFLALLGATVAFPIDDDKIVGGYTCCEENSVYQVSLNAGYHFCGSLSDQWVVS 60
 QY 61 AAHCYKSRIOVRLGEYNIIDVLEGNEOFINSKVIHNPNTNSWILDNDIMLIKLSPPAVLN 120
 DB 61 AAHCYKSRIOVRLGEYNIIDVLEGNEOFINSKVIHNPNTNSWILDNDIMLIKLSPPAVLN 120
 QY 121 ARVATISLPACAPGTCQCLISGWNTLSGTYNPELLOCLDAPILTOACEASYPGQIT 180
 DB 121 SRVSATLAPKSCPAAGTCQCLISGWNTLSGTYNPELLOCLDAPILTOACEASYPGQIT 180
 QY 181 ENMICAGFLEGGKDCOGSDGSGPVCVNGELQIVSGYCAQKPKGVYTKVCFNFDWIO 240
 DB 181 SNMCLGYMEGKDCOGSDGSGPVCVNGELQIVSGYCAQKPKGVYTKVCFNFDWIO 240
 QY 241 STIAAN 246
 DB 241 QTIAAN 246
 RESULT 13
 AAB98503
 ID AAB98503 standard; Protein; 225 AA.
 AC AAB98503;
 XX
 XX 03-AUG-2001 (first entry)
 DT
 DE Human trypsin serine protease catalytic domain.
 XX
 XX Human; TADG-15; cytostatic; vaccine; ovarian tumour; cancer; trypsin;
 KW tumour antigen-derived gene 15; serine protease.
 XX
 OS Homo sapiens.
 XX
 XX WO200129056-A1.
 XX
 XX 26-APR-2001.
 XX
 XX 20-OCT-2000; 2000WO-US29095.
 XX
 XX 20-OCT-1999; 99US-0421213.
 XX
 XX (UYAR-) UNIV ARKANSAS.
 XX
 XX O'Brien TJ, Tanimoto H;
 XX
 XX WPI; 2001-381031/40.
 XX
 XX Novel extracellular serine protease, termed tumor antigen-derived gene
 XX 15 protein overexpressed in carcinomas and DNA encoding it, for
 XX diagnosis, treatment, prevention of cancer, particularly breast,
 XX

PT ovarian cancer -
 XX
 PS Example 10; Fig 1; 130pp; English.
 XX
 CC The present invention relates to human tumour antigen-derived gene 15
 CC (TAGD-15) protein and coding sequence (see AAH23601 and AAB98500).
 CC TAGD-15 is an extracellular serine protease. It was found that TAGD-15 is
 CC over-expressed in ovarian tumours. TAGD-15 protein or its fragments of
 CC 9-20 residues that lack TAGD-15 protease activity are useful for
 CC vaccinating an individual against TAGD-15, having, suspected of having or
 CC at risk of getting cancer. Furthermore, the TAGD-15 gene can be used as a
 CC diagnostic or therapeutic target in cancer. The present sequence was used
 CC in a sequence homology alignment with the catalytic domain of TAGD-15.
 XX
 XX Sequence 225 AA;
 Query Match 75.5%; Score 1002; DB 22; Length 225;
 Best Local Similarity 80.9%; Pred. No. 4.5e-84;
 Matches 182; Conservative 23; Mismatches 20; Indels 0; Gaps 0;
 Db 23 KIVGYTCSENSVPYQVSLNAGYHFCGSLISDQWVYSAACHYKSRIOVRLGEYNIDVLE 82
 1 KIVGYTCSENSVPYQVSLNAGYHFCGSLISDQWVYSAACHYKSRIOVRLGEYNIEVLE 60
 Qy 83 GNEQFINSAKVIIRHPNYNSWILDNDIMLIKSSPAVLNARVATISLPACAPGTQCLIS 142
 Db 61 GNEQFINAAKIIIRHPQYDRKTLNDIMLIKSSRAVINARVSTISLPTAPPATGTCKLIS 120
 Qy 143 GWNTLSSGTNYPELLOCLDAPILTOACEASYPGQITENNICAGFLEGGKDCQGDGG 202
 Db 121 GWNTASSGADYPDELQCLDAPVLSQAKCEASYPGKITSNMFCVGFLEGGKDCQGDGG 180
 Qy 203 PIVCNGELOGIVSWGCGCAQKNKPGVYTKVCFNFDWIOSTIAANS 247
 Db 181 PIVCNGOLQGVVSWGCGCAQKNKPGVYTKVYNYVKWIKNTIAANS 225
 RESULT 14
 AAY31160
 ID AAY31160 standard; protein; 224 AA.
 XX
 AC AAY31160;
 XX
 DT 26-OCT-1999 (first entry)
 XX
 DE Human trypsin serine protease protein domain.
 XX
 -XW Macrophage stimulating protein; MSP; human; modulator; proliferation;
 v differentiation; intestinal epithelium; colon crypt; treatment; cancer;
 W haematopoietic disorder; megakaryocyte deficiency; gastrointestinal;
 KW chemotherapeutic agent; gut toxicity; serine protease; trypsin.
 XX
 OS Homo sapiens.
 XX
 PN US5948892-A.
 XX
 PD 07-SEP-1999.
 XX
 PF 16-DEC-1996; 96US-0766982.
 XX
 PR 16-DEC-1996; 96US-0766982.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Wahl RC;
 XX
 DR WPI; 1999-517975/43.
 XX
 XX Analogues of macrophage stimulating protein for treating
 PT gastrointestinal or haematopoietic disorders
 XX
 PS Example 2; Column 27-30; 23pp; English.
 XX

CC This invention describes a novel purified and isolated analogue of mature
 CC macrophage stimulating protein (MSP) having at least one unpaired
 CC cysteine residue substituted with another amino acid which modulates the
 CC proliferation or differentiation of the intestinal epithelium. The
 CC product of the invention binds to RON (a cell membrane protein tyrosine
 CC kinase which is a member of the c-met family) to promote the formation of
 CC colon crypts. MSP analogues are useful for the treatment of conditions
 CC requiring the administration of MSP, such conditions include
 CC haematopoietic disorders such as those involving a deficiency of
 CC megakaryocytes and gastrointestinal disorders such as ulcerative colitis,
 CC Crohn's disease and infections. The MSP analogues are useful for
 CC maintaining and repairing the epithelial lining in the treatment of
 CC cancer, where the aggressive use of chemotherapeutic agents or the use of
 CC whole body radiation may lead to gut toxicity. The MSP analogues, which
 CC have a higher activity than normal human MSP are effective at smaller
 CC dosages, or optionally, they may be administered less frequently than
 CC human MSP. This sequence represents a human trypsin serine protease
 CC domain which is used in a description of the method of the invention.
 XX
 XX Sequence 224 AA;
 Query Match 75.1%; Score 997; DB 20; Length 224;
 Best Local Similarity 80.8%; Pred. No. 1.3e-83;
 Matches 181; Conservative 23; Mismatches 20; Indels 0; Gaps 0;
 Qy 24 IVGYTCSENSVPYQVSLNAGYHFCGSLISDQWVYSAACHYKSRIOVRLGEYNIDVLE 83
 Db 1 IVGYTCSENSVPYQVSLNAGYHFCGSLISDQWVYSAACHYKSRIOVRLGEYNIEVLE 60
 Qy 84 NEQFINSAKVIIRHPNYNSWILDNDIMLIKSSPAVLNARVATISLPACAPGTQCLIS 143
 Db 61 NEQFINAAKIIIRHPQYDRKTLNDIMLIKSSRAVINARVSTISLPTAPPATGTCKLIS 120
 Qy 144 GWNTLSSGTNYPELLOCLDAPILTOACEASYPGQITENNICAGFLEGGKDCQGDGG 203
 Db 121 GWNTASSGADYPDELQCLDAPVLSQAKCEASYPGKITSNMFCVGFLEGGKDCQGDGG 180
 Qy 204 VVNCNGELOGIVSWGCGCAQKNKPGVYTKVCFNFDWIOSTIAANS 247
 Db 181 VVNCNGOLQGVVSWGCGCAQKNKPGVYTKVYNYVKWIKNTIAANS 224
 RESULT 15
 AAY77494
 ID AAY77494 standard; Protein; 229 AA.
 XX
 AC AAY77494;
 XX
 DT 05-JUN-2000 (first entry)
 XX
 DE Bovine trypsinogen.
 XX
 DE Bovine trypsinogen.
 KW Bovine; trypsinogen; heterologous protease; expression; transgenic plant;
 KW large-scale production.
 OS Bos taurus.
 XX
 FH Key Location/Qualifiers
 FT Protein 7..229
 FT /note= "Bovine trypsinogen"
 XX
 PN WO200005384-A1.
 XX
 PD 03-FEB-2000.
 XX
 PF 20-JUL-1999; 99WO-US12835.
 XX
 PR 22-JUL-1998; 98US-0120582.
 XX
 PA (PROD-) PRODIGENE INC.
 XX
 PI Howard JA, Hood E;
 XX

DR WPI; 2000-195105/17.
 DR N-PSDB; AA292585.
 XX
 PT Transgenic plant containing DNA encoding a heterologous protein, used
 PT for the commercial production of protease
 XX
 PS
 XX Example 1; Page 28-29; 32pp; English.
 XX
 CC The invention relates to a transgenic plant that comprises DNA encoding
 CC an active or inactive form of a heterologous protease (e.g., trypsin)
 CC The nucleotide sequence encoding the protease is operably linked to a
 CC promoter to effect expression of the protease in the plant. The
 CC transgenic plant is used in a method of producing protease in commercial
 CC quantities, comprising the extraction of the protease from the plant
 CC biomass. The current source of trypsin is the organs of animals. One
 CC of the main difficulties with this source is that there is considerable
 CC contamination by other proteases. Chymotrypsin is one of the additional
 CC in the contaminants, and may cleave the product in an undesired manner.
 CC The method of the invention overcomes the problem of contamination by
 CC other additional proteases. By directing the expression of the protease
 CC to the seed of the plant, the problem of expressing the protease and
 CC having it digest the proteins in the plant, resulting in cell death is
 CC overcome. The present sequence represents bovine trypsinogen. cDNA
 CC (Genbank accession #P00760) encoding this protease was used in the
 CC construction of a plant expression plasmid in the exemplifications of
 CC the invention.
 XX
 SQ Sequence 229 AA;
 Query Match 73.7%; Score 978; DB 21; Length 229;
 Best Local Similarity 77.2%; Pred. No. 7.4e-82;
 Matches 176; Conservative 29; Mismatches 23; Indels 0; Gaps 0;
 QY 19 DDDDKIVGGYTCENSVQVSLNAGYHFCGGLISDQWVSAACHYKSRIOVRLGEYNI 78
 Db 2 DDDDKIVGGYTCANTVPYQVSLNSGYHFCGGLINSQWVSAACHYKSGIOVRLGEDNI 61
 QY 79 DVLEGEQFINSKAVIRHPNYSWILNDIMLIKLSPAVLNARVATISLPACAPGTQ 138
 Db 62 NVVEGNEQFISAKSIVHPNSYNLTNNDIMLIKLSAASLNSRVASISLPTSCASAGTQ 121
 QY 139 CLISGWNLTSSGTNYPELLOCLDAPILTAQCEASYPGQITENMICAGLEGKDSQCG 198
 Db 122 CLISGWNKTSGETSYPDVLKCLKAPILSDSCKSAYPGQITSNMFCAGLEGKDSQCG 181
 QY 199 DSGGPVVCNGELQIVSWGYGCAQKNKPGVYTKVCFVDWIQSTIAAN 246
 Db 182 DSGGPVVCCKLOGIVSWGSGCAQKNKPGVYTKVCFVSWIKOTIASN 229

Search completed: June 20, 2003, 15:19:52
 Job time : 82 secs

Db. 63 AHCKYSRIQVRLGHNIDVLENEQFINAAKIIHPNENGNTLNDIMLIKLSPPATLNS 122
 QY 122 RVATISLPRAACAAPCTOCLISGWNTLSSTGNYPELLOCLDAPILTOAQCEASYPQGIT 181
 Db 123 RVATISLPRAACAAGTECLISGWNTLSSTGNYPELLOCLDAPILTOAQCEASYPQGIT 182
 QY 182 NMICAGFLEGGKDCSQDGGPVVNCNGELQISVSWGYGCAQKNKPGVYTKVCNFDVWIO 241
 Db 183 NMICVGFLEGGKDCSQDGGPVVNCNGELQISVSWGYGCAQKNKPGVYTKVCNFDVWIO 242
 QY 242 TIAAN 246
 Db 243 TIAAN 247

RESULT 2

US-08-978-404B-44
 ; Sequence 44, Application US/08978404B
 ; Patent No. 5968782

GENERAL INFORMATION:

APPLICANT: Stevens, Richard L.
 TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
 TITLE OF INVENTION: FIBRINOGEN
 NUMBER OF SEQUENCES: 74
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: U.S.A.
 ZIP: 02210-2211

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/978.404B
 FILING DATE: 25-NOV-97
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/032,354
 FILING DATE: 04-DEC-1996

ATTORNEY/AGENT INFORMATION:

NAME: Plumer, Elizabeth R.
 REGISTRATION NUMBER: 36,637
 REFERENCE/DOCKET NUMBER: B0801/7090
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441
 TELEX:

INFORMATION FOR SEQ ID NO: 44:

SEQUENCE CHARACTERISTICS:
 LENGTH: 246 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: No. 5968782e

US-08-978-404B-44

Query Match 83.0%; Score 1101; DB 2; Length 246;
 Best Local Similarity 80.5%; Pred. No. 1.5e-106;
 Matches 198; Conservative 26; Mismatches 22; Indels 0; Gaps 0;

QY 1 MNPLILALFLGAATPTDDDKIVGGYTCBENSVPYQVSLNAGYHFCGSLISQWVVS 60
 Db 1 MSALLILALVGAAPFLEDDDKIVGGYTCBENSVPYQVSLNAGYHFCGSLISQWVVS 60
 QY 61 AAHYCKRSIQVRLGEYNDVLENEQFINAAKIIHPNENGNTLNDIMLIKLSPPAVLN 120
 Db 61 AAHYCKRSIQVRLGEYNDVLENEQFINAAKIIHPNENGNTLNDIMLIKLSPPAVLN 120

QY 121 ARVATISLPRAACAAPCTOCLISGWNTLSSTGNYPELLOCLDAPILTOAQCEASYPQGIT 180
 Db 121 ARVAPVALPSACAPAGTQCLISGWNTLSNGVNNPDLLQCDAPVLSQADCEAAYPGEIT 180
 QY 181 ENMICAGFLEGGKDCSQDGGPVVNCNGELQISVSWGYGCAQKNKPGVYTKVCNFDVWIO 240
 Db 181 SSMICVGFLEGGKDCSQDGGPVVNCNGELQISVSWGYGCAQKNKPGVYTKVCNFDVWIO 240
 QY 241 STIAAN 246
 Db 241 DTIAAN 246

RESULT 3

US-08-467-155A-7
 ; Sequence 7, Application US/08467155A
 ; Patent No. 5736377

GENERAL INFORMATION:

APPLICANT: Band, Vmla
 TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED
 TITLE OF INVENTION: MOLECULES AND METHODS
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/467,155A
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:

NAME: Clark, Paul T.
 REGISTRATION NUMBER: 30,162
 REFERENCE/DOCKET NUMBER: 00398/100001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/542-5070
 TELEFAX: 617/542-8906
 TELEX: 200154

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:
 LENGTH: 281 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: linear
 MOLECULE TYPE: protein

US-08-467-155A-7

Query Match 79.8%; Score 1059.5; DB 1; Length 281;
 Best Local Similarity 72.2%; Pred. No. 3.8e-102;
 Matches 203; Conservative 20; Mismatches 23; Indels 35; Gaps 5;

QY 1 MNPLILALFLGAATPTDDDKIVGGY-----TCEBNSVPYQVSLNAGYHFCGSLI 53
 Db 1 MSALLILALVGAAPFVDDDDKIVGYMMTRYARTCRESSVPYQVSLNAGYHFCGSLI 60
 QY 54 SDQWVSAHCKYSRIQVRLGEYN-----IDVLENEQFINSAKVIHPNNTNSWTLN 106
 Db 61 NDQWVSAHCKYSRIQVRLGEHNMTRYARINVLNEQFVDSAKIIRHPNNTNSWTLN 120
 QY 107 DIMLIKLSPPAVL-----NARVATISLPRAACPGTCLISGWNTLSSTGNYPELLO 159
 Db 121 DIMLIKLSPPAVLMTMTRYARNARVASVPLPSCAPAGTQCLISGWNTLSNGVNNPDLLQ 180
 QY 160 CIDAPILTOA-----OCEASYPGQITENMTCAGFLEGGKDCSQDGGPVVNCNGELQ 212
 Db 160 CIDAPILTOA-----OCEASYPGQITENMTCAGFLEGGKDCSQDGGPVVNCNGELQ 212

Db 181 CVDAPVLPQAMMTRYARDCEASYPGDIITNNMICVGFLEGGKDCSCQDSGGPVVNCNGLAQ 240
QY 213 IVSWGYP-----GCAQKNKPGVYTKVCNFDVWQIQTIAAN 246
Db 241 IVSWGIMTRYARGCAQPDAPGVYTKVCNFDVWQIQTIAADN 281

RESULT 4
US-08-628-198-7
; Sequence 7, Application US/08628198
; Patent No. 5843694
; GENERAL INFORMATION:
; APPLICANT: Band, Vimla
; TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED
; TITLE OF INVENTION: MOLECULES AND METHODS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/628,198
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/467,155
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00398/100002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-628-198-7

Query Match 79.8%; Score 1059.5; DB 2; Length 281;
Best Local Similarity 72.2%; Pred. No. 3.8e-102;
Matches 203; Conservative 20; Mismatches 23; Indels 35; Gaps 5;
QY 1 MNPILLILAFGAAVATPTDDDKIVGGY-----TCEENSYPYQVSLNAGYHFCGGSII 53
Db 1 MSALLILALVGAAVAFVDDDDKIVGGYMMTRYARTCRESSYPYQVSLNAGYHFCGGSII 60
QY 54 SDQWVYSAACHYKSRIOVRLGEYN-----IDVLENEQFINSKVIIRHPNYSWILDN 106
Db 61 NDQWVYSAACHYKRIQVRLGEHNMTRYARINVLGEQFVDSAKIIRHPNYSWILDN 120
QY 107 DIMLIKSSPAVL-----NARVATISLPACAPGTQCLISGHWNTLSSTNTYPELLQ 159
Db 121 DIMLIKASPTLMMTRYARNARVASVPLPSSCAPAGTQCLISGHWNTLSGNNPDLQ 180
QY 160 CLDAPILTQA-----QCEASYPGOITENNIMICAGFLEGGKDCSCQDSGGPVVNCNGLAQ 212
Db 181 CVDAPVLPQAMMTRYARDCEASYPGDIITNNMICVGFLEGGKDCSCQDSGGPVVNCNGLAQ 240
QY 213 IVSWGYP-----GCAQKNKPGVYTKVCNFDVWQIQTIAAN 246

Db 241 IVSWGIMTRYARGCAQPDAPGVYTKVCNFDVWQIQTIAADN 281
RESULT 5
US-09-201-038-7
; Sequence 7, Application US/09201038
; Patent No. 6153387
; GENERAL INFORMATION:
; APPLICANT: Band, Vimla
; TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED
; TITLE OF INVENTION: MOLECULES AND METHODS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/201,038
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/628,198
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00398/100002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-201-038-7

Query Match 79.8%; Score 1059.5; DB 4; Length 281;
Best Local Similarity 72.2%; Pred. No. 3.8e-102;
Matches 203; Conservative 20; Mismatches 23; Indels 35; Gaps 5;
QY 1 MNPILLILAFGAAVATPTDDDKIVGGY-----TCEENSYPYQVSLNAGYHFCGGSII 53
Db 1 MSALLILALVGAAVAFVDDDDKIVGGYMMTRYARTCRESSYPYQVSLNAGYHFCGGSII 60
QY 54 SDQWVYSAACHYKSRIOVRLGEYN-----IDVLENEQFINSKVIIRHPNYSWILDN 106
Db 61 NDQWVYSAACHYKRIQVRLGEHNMTRYARINVLGEQFVDSAKIIRHPNYSWILDN 120
QY 107 DIMLIKSSPAVL-----NARVATISLPACAPGTQCLISGHWNTLSSTNTYPELLQ 159
Db 121 DIMLIKASPTLMMTRYARNARVASVPLPSSCAPAGTQCLISGHWNTLSGNNPDLQ 180
QY 160 CLDAPILTQA-----QCEASYPGOITENNIMICAGFLEGGKDCSCQDSGGPVVNCNGLAQ 212
Db 181 CVDAPVLPQAMMTRYARDCEASYPGDIITNNMICVGFLEGGKDCSCQDSGGPVVNCNGLAQ 240
QY 213 IVSWGYP-----GCAQKNKPGVYTKVCNFDVWQIQTIAAN 246
Db 241 IVSWGIMTRYARGCAQPDAPGVYTKVCNFDVWQIQTIAADN 281

RESULT 7
US-08-944-483-34

US-09-027-337-5
 ; Sequence 5, Application US/09027337B
 ; Patent No. 5972616
 ; GENERAL INFORMATION:
 ; APPLICANT: O'Brien, Timothy J.
 ; APPLICANT: Tanimoto, Hiroto
 ; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease Overexpressed in
 ; FILE REFERENCE: D6064
 ; CURRENT APPLICATION NUMBER: US/09/027.337B
 ; CURRENT FILING DATE: 1998-02-20
 ; NUMBER OF SEQ ID NOS: 13
 ; SEQ ID NO 5
 ; LENGTH: 225
 ; TYPE: PRT
 ; ORGANISM: Unknown
 ; FEATURE:
 ; OTHER INFORMATION: Serine protease catalytic domain of trypsin (Try)
 ; OTHER INFORMATION: homologous to similar domain in TADG-15
 ; S-09-027-337-5

Query Match 75.5%; Score 1002; DB 2; Length 225;
 Best Local Similarity 80.9%; Pred. No. 2.6e-96;
 Matches 182; Conservative 23; Mismatches 20; Indels 0; Gaps 0;
 QY 23 KIVGYTCSENSVPYQVSLNAGYHFCGSLSDQVYSAACHYKSRIOVRLGEYNIDVLE 82
 DB 1 KIVGYNCEENSVPYQVSLNAGYHFCGSLSDQVYSAACHYKSRIOVRLGEYNIEVLE 60
 QY 83 GNEQFINSKAVIRHPNYSWILNDIMLIKSSPAVLNARVATISLPACAPGTQCLIS 142
 DB 61 GNEQFINAAKIIIRHPQYDRKTLNNDIMLIKSSRAVINARVSTISLPTAPPATGTCKLIS 120
 QY 143 GWNTLSSGNTYPELLQCLDAPILTOAQCEASYPGQITENNICAGFLEGGKDCSGDGG 202
 DB 121 GWNTASSGADYPDELQCLDAPVLSQAKCEASYPGKITSNMFCVGLGGKDCSGDGG 180
 QY 203 PIVCNGELQGVSWGDCGCAQKNKPGVYTKVCFVDMQSTIAANS 247
 DB 181 PIVCNGQLGVSWGDCGCAQKNKPGVYTKVYNYVKWIKNTIAANS 225

RESULT 9
 US-09-644-600-5
 ; Sequence 5, Application US/09644600
 ; Patent No. 6451500
 ; GENERAL INFORMATION:
 ; APPLICANT: O'Brien, Timothy J.
 ; APPLICANT: Tanimoto, Hiroto
 ; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease
 ; FILE REFERENCE: D6064CIP/D
 ; CURRENT APPLICATION NUMBER: US/09/644.600
 ; CURRENT FILING DATE: 2000-08-23
 ; PRIOR APPLICATION NUMBER: 09/421,213
 ; PRIOR FILING DATE: 1999-10-20
 ; PRIOR APPLICATION NUMBER: 09/027,337
 ; PRIOR FILING DATE: 1998-02-20
 ; NUMBER OF SEQ ID NOS: 98
 ; SEQ ID NO 5
 ; LENGTH: 225
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Trypsin
 ; US-09-644-600-5

Query Match 75.5%; Score 1002; DB 4; Length 225;
 Best Local Similarity 80.9%; Pred. No. 2.6e-96;
 Matches 182; Conservative 23; Mismatches 20; Indels 0; Gaps 0;
 QY 23 KIVGYTCSENSVPYQVSLNAGYHFCGSLSDQVYSAACHYKSRIOVRLGEYNIDVLE 82
 DB 1 KIVGYNCEENSVPYQVSLNAGYHFCGSLSDQVYSAACHYKSRIOVRLGEYNIEVLE 60

DB 1 KIVGYNCEENSVPYQVSLNAGYHFCGSLSDQVYSAACHYKSRIOVRLGEYNIEVLE 60
 QY 83 GNEQFINSKAVIRHPNYSWILNDIMLIKSSPAVLNARVATISLPACAPGTQCLIS 142
 DB 61 GNEQFINAAKIIIRHPQYDRKTLNNDIMLIKSSRAVINARVSTISLPTAPPATGTCKLIS 120
 QY 143 GWNTLSSGNTYPELLQCLDAPILTOAQCEASYPGQITENNICAGFLEGGKDCSGDGG 202
 DB 121 GWNTASSGADYPDELQCLDAPVLSQAKCEASYPGKITSNMFCVGLGGKDCSGDGG 180
 QY 203 PIVCNGELQGVSWGDCGCAQKNKPGVYTKVCFVDMQSTIAANS 247
 DB 181 PIVCNGQLGVSWGDCGCAQKNKPGVYTKVYNYVKWIKNTIAANS 225

RESULT 10
 US-08-766-982-13
 ; Sequence 13, Application US/08766982
 ; Patent No. 5948892
 ; GENERAL INFORMATION:
 ; APPLICANT: Wahl, Robert C.
 ; TITLE OF INVENTION: Analogs of Macrophage Stimulating
 ; TITLE OF INVENTION: Protein
 ; NUMBER OF SEQUENCES: 13
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Amgen Inc.
 ; STREET: 1840 Behavilland Drive
 ; CITY: Thousand Oaks
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 91320-1789
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/766,982
 ; FILING DATE:
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Winter, Robert B.
 ; REFERENCE/DOCKET NUMBER: A-441
 ; INFORMATION FOR SEQ ID NO: 13:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 224 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-766-982-13

Query Match 75.1%; Score 997; DB 2; Length 224;
 Best Local Similarity 80.8%; Pred. No. 8.7e-96;
 Matches 181; Conservative 23; Mismatches 20; Indels 0; Gaps 0;
 QY 24 IVGGYTCEENSVPYQVSLNAGYHFCGSLSDQVYSAACHYKSRIOVRLGEYNIDVLE 83
 DB 1 IVGGYNCEENSVPYQVSLNAGYHFCGSLSDQVYSAACHYKSRIOVRLGEYNIEVLE 60
 QY 84 NEQFINSKAVIRHPNYSWILNDIMLIKSSPAVLNARVATISLPACAPGTQCLIS 143
 DB 61 NEQFINAAKIIIRHPQYDRKTLNNDIMLIKSSRAVINARVSTISLPTAPPATGTCKLIS 120
 QY 144 GWNTLSSGNTYPELLQCLDAPILTOAQCEASYPGQITENNICAGFLEGGKDCSGDGG 203
 DB 121 GWNTASSGADYPDELQCLDAPVLSQAKCEASYPGKITSNMFCVGLGGKDCSGDGG 180
 QY 204 VVNCNGELQGVSWGDCGCAQKNKPGVYTKVCFVDMQSTIAANS 247
 DB 181 VVNCNGQLGVSWGDCGCAQKNKPGVYTKVYNYVKWIKNTIAANS 224

RESULT 11
US-08-944-483-36
; Sequence 36, Application US/08944483
; Patent No. 6232456
; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STEWART, KENT D.
; APPLICANT: STROUPE, STEVEN D.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
; TITLE OF INVENTION: OF THE PROSTATE
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,483
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6183.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 224 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6232456e
; US-08-944-483-36

Query Match 75.1%; Score 997; DB 4; Length 224;
Best Local Similarity 80.8%; Pred. No. 8.7e-96;
Matches 181; Conservative 23; Mismatches 20; Indels 0; Gaps 0;

Qy 24 IVGGYTCENSVYQVSLNAGYHFCGSLSDQVWVSAACHYKSRQVRLGEYNIDVLEG 83
Db 1 IVGGYTCENSVYQVSLNAGYHFCGSLSDQVWVSAACHYKSRQVRLGEYNIDVLEG 60

Qy 84 NEQFINSAKVIHRPNYNSWILDNDIMLKSSPAVLNARVATISLPACAPGTQCLISG 143
Db 61 NEQFINAAKIIHRPQYDRKTLNNDIMLKSSRAVINARVSTISLPTAPPATGKCLISG 120

Qy 144 WGNLTSSGTYNPELLOCLDAPILTAQACEASYPGQITENMICAGFLEGGKDCQDGGGP 203
Db 121 WGNLTSSGTYNPELLOCLDAPILTAQACEASYPGQITENMICAGFLEGGKDCQDGGGP 180

Qy 204 VVCNGELQGVVSWGCGCAQKNKPGVYTKVNCNFVDWIOSTIAANS 247
Db 181 VVCNGELQGVVSWGCGCAQKNKPGVYTKVNCNFVDWIOSTIAANS 224

; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS

RESULT 12
US-09-296-219-13
; Sequence 13, Application US/09296219
; Patent No. 6248560
; GENERAL INFORMATION:
; APPLICANT: Wahl, Robert C.
; TITLE OF INVENTION: Analogs of Macrophage Stimulating
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/296,219
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-441
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 224 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-296-219-13

Query Match 75.1%; Score 997; DB 4; Length 224;
Best Local Similarity 80.8%; Pred. No. 8.7e-96;
Matches 181; Conservative 23; Mismatches 20; Indels 0; Gaps 0;

Qy 24 IVGGYTCENSVYQVSLNAGYHFCGSLSDQVWVSAACHYKSRQVRLGEYNIDVLEG 83
Db 1 IVGGYTCENSVYQVSLNAGYHFCGSLSDQVWVSAACHYKSRQVRLGEYNIDVLEG 60

Qy 84 NEQFINSAKVIHRPNYNSWILDNDIMLKSSPAVLNARVATISLPACAPGTQCLISG 143
Db 61 NEQFINAAKIIHRPQYDRKTLNNDIMLKSSRAVINARVSTISLPTAPPATGKCLISG 120

Qy 144 WGNLTSSGTYNPELLOCLDAPILTAQACEASYPGQITENMICAGFLEGGKDCQDGGGP 203
Db 121 WGNLTSSGTYNPELLOCLDAPILTAQACEASYPGQITENMICAGFLEGGKDCQDGGGP 180

Qy 204 VVCNGELQGVVSWGCGCAQKNKPGVYTKVNCNFVDWIOSTIAANS 247
Db 181 VVCNGELQGVVSWGCGCAQKNKPGVYTKVNCNFVDWIOSTIAANS 224

; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS

;; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
;; TITLE OF INVENTION: OF THE PROSTATE
;; NUMBER OF SEQUENCES: 76
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Abbott Laboratories
;; STREET: 100 Abbott Park Road
;; CITY: Abbott Park
;; STATE: IL
;; COUNTRY: USA
;; ZIP: 60064-3500
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/944,483
;; FILING DATE:
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Becker, Cheryl L.
;; REGISTRATION NUMBER: 35,441
;; REFERENCE/DOCKET NUMBER: 6183.US.01
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 847/935-1729
;; TELEFAX: 847/938-2623
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 35:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 224 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: No. 6232456e
US-08-944-483-35

Query Match 74.9%; Score 994; DB 4; Length 224;
Best Local Similarity 80.4%; Pred. No. 1.8e-95;
Matches 180; Conservative 24; Mismatches 20; Indels 0; Gaps 0;

Qy 24 IVGGYTCEENSVPTQVSLNAGYHFCGGLSDQVWVSAHCKYSRIQVRLGEYNTDVLG 83
Db 1 IVGGYTCEENSLPYQVSLNSGSHFCGGLSDQVWVSAHCKYKTRIQVRLGEHNTKVLG 60
84 NEQFINSKAVTRHPNYSWILDNDIMLIKSSPAVLNARVATISLPACAPGTCLISG 143
61 NEQFINAKIIRHPKYNRDLNDIMLIKSSPAVLNARVATISLPTAPPAAGTECLISG 120
144 WGNLTSSGTNPPELLQCLDAPILTAQACEASYPGQITENNICAGFLEGKDSQGDSSGP 203
121 WGNLTSGADYPDELKCLDAPVLTQACEKASYPGKITNSMFCVGLGKDSQGDSSGP 180
Qy 204 VVCNGELQGVSWGSGCAQKNKPGVYTKVCFNVDWQIQTIAANS 247
Db 181 VVCNGQLQGVVSWGHGCAKKNRPGVYTKVYVVDWIKDTIAANS 224

RESULT 14
US-09-120-582-2
; Sequence 2, Application US/09120582
; Patent No. 6087558
; GENERAL INFORMATION:
; APPLICANT: Howard, John A.
; APPLICANT: Hood, Elizabeth
; TITLE OF INVENTION: COMMERCIAL PRODUCTION OF PROTEASES IN PLANTS
; FILE REFERENCE: 10014
; CURRENT APPLICATION NUMBER: US/09/120,582
; CURRENT FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.0

;; SEQ ID NO 2
;; LENGTH: 229
;; TYPE: PRT
;; ORGANISM: Bos taurus
US-09-120-582-2

Query Match 73.7%; Score 978; DB 3; Length 229;
Best Local Similarity 77.2%; Pred. No. 8.5e-94;
Matches 176; Conservative 29; Mismatches 23; Indels 0; Gaps 0;

Qy 19 DDDKIVGGYTCEENSVPTQVSLNAGYHFCGGLSDQVWVSAHCKYSRIQVRLGEYNI 78
Db 2 DDDKIVGGYTCCGANTPYQVSLNSGSHFCGGLSDQVWVSAHCKYKSGIQVRLGSDNI 61
Qy 79 DYLEGNEQFINSKAVTRHPNYSWILDNDIMLIKSSPAVLNARVATISLPACAPGTQ 138
Db 62 NVVEGNEQFISASKSIVHPSYNSNTLNDIMLIKSSPAVLNARVATISLPTSCASAGTQ 121
Qy 139 CLISGNGNTLSSGTNPPELLQCLDAPILTAQACEASYPGQITENNICAGFLEGKDSQGD 198
Db 122 CLISGNGNTKSSGTSYPTDVLKCLPILSDSSCKSAYPGQITNSMFCAGYLEGKDSQGD 181
Qy 199 DSGGPVVCNGELQGVSWGSGCAQKNKPGVYTKVCFNVDWQIQTIAAN 246
Db 182 DSGGPVVCNGELQGVSWGSGCAQKNKPGVYTKVCFNVDWQIQTIAAN 229

RESULT 15
US-08-467-155A-8
; Sequence 8, Application US/08467155A
; Patent No. 5736377
; GENERAL INFORMATION:
; APPLICANT: Band, Vimla
; TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED
; TITLE OF INVENTION: MOLECULES AND METHODS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,155A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00398/100001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 299 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-467-155A-8

Query Match 71.6%; Score 950.5; DB 1; Length 299;
Best Local Similarity 68.2%; Pred. No. 8.9e-91;
Matches 187; Conservative 24; Mismatches 22; Indels 41; Gaps 6;

```

QY 14 VATPTDDDDKIVGY-----TCENSVPYQVSLNAGYHFCGSLISDOMVYSAHCY 65
Db 27 VAVPFDDDDKIVGGYHSTRYIVBTC-ENSLPYQVSLNSGSHFCGSLISEQWVYSAHCY 85
QY 66 KSRIOVRLGEVN-----IDVLEGNQOFINSKVIHHPNYNSWILDNDIMLIKLSPPA 117
Db 86 KTRIOVRLGEHNSHSTRYIVBIVKIVLEGNQOFINAAKIIHHPKYNRDTLNDNDIMLIKLSPPA 145
QY 118 VL-----NARVATISLPRACAPGTQCLISGWNTLSSGTNYPELLOCLDAPILTOA 169
Db 146 VIHSTRYIVBNARVSTISLPTAPPAAGTECLISGWNTLSFGADYPDELKCLDAPVLTOA 205
QY 170 -----QCEASYPGOITENMICAGFLEGGKDCQDSDSGGPPVVCNGELQIVSWGY--- 218
Db 206 HSTRYIVBECKASYPGXITNSMFCVGFLEGGKDCQDSDSGGPPVVCNGELQIVSWGHST 265
QY 219 -----GCAOKNKGVTYKVCNFVDWIOSTIAANS 247
Db 266 RYIVBGCAWKNRPGVYTKVYNVDWIKDTIAANS 299

```

-arch completed: June 20, 2003, 15:23:14
Job time : 34 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 15:17:26 ; Search time 40 seconds
(without alignments)
593.630 Million cell updates/sec

Title: US-09-762-277A-1

Perfect score: 1327

Sequence: 1 MNPLILAFILGAAVATPTDD.....VYTKVCFNFDWIQSTIAANS 247

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

otal number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: piri:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1327	100.0	247	1 TRDG	trypsin (EC 3.4.21
2	1153	86.9	247	2 S13813	trypsin (EC 3.4.21
3	1127	84.9	246	2 B25528	trypsin (EC 3.4.21
4	1107	83.4	246	1 TRRT2	trypsin (EC 3.4.21
5	1101	83.0	246	1 TRRT1	trypsin (EC 3.4.21
6	1098.5	82.8	247	2 A27547	trypsin (EC 3.4.21
7	1097	82.7	247	1 B25852	trypsin (EC 3.4.21
8	1085	81.8	247	1 A25852	trypsin (EC 3.4.21
9	1071	80.7	231	1 TRPQTR	trypsin (EC 3.4.21
10	1058	79.7	247	1 S12764	trypsin (EC 3.4.21
11	1036	78.1	304	2 S33496	trypsin (EC 3.4.21
12	1026	77.3	246	1 TRDGC	trypsin (EC 3.4.21
13	1020.5	76.9	259	2 I38363	trypsin (EC 3.4.21
14	992.5	74.8	243	2 A35871	trypsin (EC 3.4.21
15	978	73.7	229	1 TRBQTR	trypsin (EC 3.4.21
16	966.5	72.8	248	2 S55066	trypsin (EC 3.4.21
17	960.5	72.4	247	2 S05494	trypsin (EC 3.4.21
18	948.5	71.5	238	2 S31779	trypsin (EC 3.4.21
19	943	71.1	248	2 S55067	trypsin (EC 3.4.21
20	905	68.2	246	2 J01472	trypsin (EC 3.4.21
21	896	67.5	246	2 J01471	trypsin (EC 3.4.21
22	877	66.1	242	2 S31775	trypsin (EC 3.4.21
23	877	66.1	242	2 S31776	trypsin (EC 3.4.21
24	867	65.3	231	2 S31778	trypsin (EC 3.4.21
25	855.5	64.5	229	1 TRDES	trypsin (EC 3.4.21
26	854	64.4	242	2 S49489	trypsin (EC 3.4.21
27	834.5	62.9	241	2 S39048	trypsin (EC 3.4.21
28	820.5	61.8	240	2 S39047	trypsin (EC 3.4.21
29	690.5	52.0	250	2 T01779	trypsin (EC 3.4.21

30 649.5 48.9 250 2 S31384 trypsin (EC 3.4.21
31 555 41.8 260 2 I06559 neuropsin - mouse
32 535 40.3 262 1 KOHU tissue kallikrein
33 534 40.2 266 2 S54146 trypsin (EC 3.4.21
34 530.5 40.0 261 2 S45303 tissue kallikrein
35 528.5 39.8 257 2 S33772 tissue kallikrein
36 527.5 39.8 253 2 A53968 serine proteinase
37 516.5 38.9 261 2 A25606 tissue kallikrein
38 508 38.3 261 2 A29586 tissue kallikrein
39 501.5 37.8 233 1 JG0169 venombin A (EC 3.4
40 500 37.7 261 2 A31136 tissue kallikrein
41 500 37.7 261 2 A34079 tissue kallikrein
42 499.5 37.6 261 1 KQMS1 tissue kallikrein
43 498 37.5 232 1 KQPG tissue kallikrein
44 498 37.5 259 2 B31136 tissue kallikrein
45 497.5 37.5 265 1 KQRTF tissue kallikrein

ALIGNMENTS

RESULT 1

TRDG

trypsin (EC 3.4.21.4) precursor, anionic - dog

N:Alternate names: cationic trypsinogen

C:Species: Canis lupus familiaris (dog)

C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999

C:Accession: A26273

R:Pinsky, S.D.; LaForge, K.S.; Scheele, G.

Mol. Cell. Biol. 5, 2669-2676, 1985

A:Title: Differential regulation of trypsinogen mRNA translation: full-length mRNA s

A:Reference number: A26273; MUID:86284628; PMID:3841794

A:Accession: A26273

A:Molecule type: mRNA

A:Residues: 1-247 <PIN>

A:Cross-references: GB:M11589; NID:g164094; PIDN:AAA30899.1; PID:g164095

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen

F:1-15/Domain: signal sequence #status predicted <SIG>

F:16-23/Domain: activation peptide #status predicted <APT>

F:24-247/Product: trypsin, anionic #status predicted <ENZ>

F:24-239/Domain: trypsin homology <TRY>

F:30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted

F:63,107,200/Active site: His, Asp, Ser #status predicted

F:75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 100.0%; Score 1327; DB 1; Length 247;

Best Local Similarity 100.0%; Pred. No. 1.2e-105;

Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNPLILAFILGAAVATPTDDDKIVGGYTCENSVPYQVSLNAGYHFCGSLISDQWVS 60

DB 1 MNPLILAFILGAAVATPTDDDKIVGGYTCENSVPYQVSLNAGYHFCGSLISDQWVS 60

QY 61 AAHCYKSIQVRLGEYNDVLENGEQFINSAKVRHPNTNSWILDNDIMLIKUSSPAVLN 120

DB 61 AAHCYKSIQVRLGEYNDVLENGEQFINSAKVRHPNTNSWILDNDIMLIKUSSPAVLN 120

QY 121 ARVATISLPRACAAPGTCTCLSGWNTLSSGNTYNPYELQCLDAPILTOQCEASYPGQIT 180

DB 121 ARVATISLPRACAAPGTCTCLSGWNTLSSGNTYNPYELQCLDAPILTOQCEASYPGQIT 180

QY 181 ENMICAGLEGKDCSQDGGPVPVNCNGELQGVSWNGYCAQKNKPGVYTKVCFNFDWIQ 240

DB 181 ENMICAGLEGKDCSQDGGPVPVNCNGELQGVSWNGYCAQKNKPGVYTKVCFNFDWIQ 240

QY 241 STIAANS 247

DB 241 STIAANS 247

RESULT 2

S13813

trypsin (EC 3.4.21.4) - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 02-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999
 C:Accession: S13813
 R:Le Huereu, I.; Wicker, C.; Guilloteau, P.; Toullec, R.; Puigserver, A.
 Eur. J. Biochem. 193, 767-773, 1990
 A:Title: Isolation and nucleotide sequence of cDNA clone for bovine pancreatic anionic
 A:Reference number: S13813; MUID:91065383; PMID:1701147
 A:Accession: S13813
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-247 <HUE>
 A:Cross-references: EMBL:X54703; NID:g829; PIDN:CAA38513.1; PID:g830
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; protein digestion; serine proteinase
 F:24-239/Domain: trypsin homology <TRY>
 F:63,107,200/Active site: His, Asp, Ser #status predicted

Query Match 86.9%; Score 1153; DB 2; Length 247;
 Best Local Similarity 83.8%; Pred. No. 8.2e-91;
 Matches 207; Conservative 22; Mismatches 18; Indels 0; Gaps 0;

QY 1 MNPLILALFGLAAGAAVATPTDDDDKIVGGYTCENSVYQVSLNAGYHFCGSLSDQWVVS 60
 Db 1 MHPLILALFGLAAGAAVATPTDDDDKIVGGYTCENSVYQVSLNAGYHFCGSLINDQWVVS 60

QY 61 AAHCYKSRIOVRLGEYNIDVLEGNQFINSKAVIRHPNYSWILNDIMLIKSSPAVLN 120
 Db 61 AAHCYQTHIOVRLGEYNIDVLEGNQFIDASKIIRHPKYSWTLNDILLIKLSTPAVLN 120

QY 121 ARVATISLPRACAAAGTQCLISGNGNTLSSTGNTYPELLQCLIDAPILTAQACEASYPGQIT 180
 Db 121 ARVSTLLPSACASAGTECLISGNGNTLSGNGYVPLLQCLVAPLLSHADCEASYPGQIT 180

QY 181 ENMICAGFLEGGKDCSGDGGPVVNCNGELQIVSWGCAQKPKGVYTKVCNFVDWIO 240
 Db 181 NMICAGFLEGGKDCSGDGGPVVNCNGELQIVSWGCAQKPKGVYTKVCNFVDWIO 240

QY 241 STIAANS 247
 Db 241 ETIAANS 247

RESULT 3
 B25528
 trypsin (EC 3.4.21.4) precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 22-Jun-1999
 C:Accession: B25528
 Stevenson, B.J.; Hagenbuechle, O.; Wellauer, P.K.
 Nucleic Acids Res. 14, 8307-8330, 1986
 A:Title: Sequence organisation and transcriptional regulation of the mouse elastase II a
 A:Reference number: A93646; MUID:87066713; PMID:3641189
 A:Accession: B25528
 A:Molecule type: mRNA
 A:Residues: 1-246 <STE>
 A:Cross-references: GB:X04574; NID:g54918; PIDN:CAA28243.1; PID:g54919
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: calcium binding; hydrolase; protein digestion; serine proteinase
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-246/Product: trypsin #status predicted <MAT>
 F:24-239/Domain: trypsin homology <TRY>
 F:30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted
 F:63,107,200/Active site: His, Asp, Ser #status predicted
 F:75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 84.9%; Score 1127; DB 2; Length 246;
 Best Local Similarity 82.5%; Pred. No. 1.3e-88;
 Matches 203; Conservative 20; Mismatches 23; Indels 0; Gaps 0;

QY 1 MNPLILALFGLAAGAAVATPTDDDDKIVGGYTCENSVYQVSLNAGYHFCGSLSDQWVVS 60
 Db 1 MSALLILALVGAAGAAVATPTDDDDKIVGGYTCRESSVYQVSLNAGYHFCGSLINDQWVVS 60

QY 61 AAHCYKSRIOVRLGEYNIDVLEGNQFINSKAVIRHPNYSWILNDIMLIKSSPAVLN 120
 Db 61 AAHCYKRIQVRLGEHINVLGNQFVDSKILFHPNYSWTLNDIMLIKASPVLN 120

QY 121 ARVATISLPRACAAAGTQCLISGNGNTLSSTGNTYPELLQCLIDAPILTAQACEASYPGQIT 180
 Db 121 ARVASVPLPSSCAPAGTQCLISGNGNTLSNGVNNPDLQCVDAPYLPQADCEASYPGKIT 180

QY 181 ENMICAGFLEGGKDCSGDGGPVVNCNGELQIVSWGCAQKPKGVYTKVCNFVDWIO 240
 Db 181 NMTCVGFLEGGKDCSGDGGPVVNCNGELQIVSWGCAQPDAPGVYTKVCNFVDWIO 240

QY 241 STIAAN 246
 Db 241 NTIADN 246

RESULT 4
 TRRT2
 trypsin (EC 3.4.21.4) II precursor - rat
 N:Alternate names: trypsinogen II
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 05-Apr-1983 #sequence_revision 30-Sep-1987 #text_change 18-Jul-1997
 C:Accession: A22657; A00949
 R:Craik, C.S.; Choo, Q.L.; Swift, G.H.; Quinto, C.; Macdonald, R.J.; Rutter, W.J.
 J. Biol. Chem. 259, 14255-14264, 1984
 A:Title: Structure of two related rat pancreatic trypsin genes.
 A:Reference number: A22657; MUID:85054880; PMID:6094547
 A:Accession: A22657
 A:Molecule type: DNA
 A:Residues: 1-246 <CRA>
 R:Macdonald, R.J.; Stary, S.J.; Swift, G.H.
 J. Biol. Chem. 257, 9724-9732, 1982
 A:Title: Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide sequences
 A:Reference number: A00948; MUID:82265624; PMID:6896710
 A:Accession: A00949
 A:Molecule type: mRNA
 A:Residues: 9-246 <MAC>
 C:Comment: The trypsin II mRNA is present in much lower quantities than the trypsin I
 C:Genetics:
 A:Introns: 14/1; 67/2
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
 F:1-15/Domain: signal sequence #status predicted <SIG>
 F:16-23/Domain: activation peptide #status predicted <APT>
 F:24-246/Product: trypsin II #status predicted <ENZ>
 F:24-239/Domain: trypsin homology <TRY>
 F:30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted
 F:63,107,200/Active site: His, Asp, Ser #status predicted
 F:75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 83.4%; Score 1107; DB 1; Length 246;
 Best Local Similarity 81.3%; Pred. No. 6.7e-87;
 Matches 200; Conservative 23; Mismatches 23; Indels 0; Gaps 0;

QY 1 MNPLILALFGLAAGAAVATPTDDDDKIVGGYTCENSVYQVSLNAGYHFCGSLSDQWVVS 60
 Db 1 MRALLFALVGAAGAAVATPTDDDDKIVGGYTCQENSVYQVSLNAGYHFCGSLINDQWVVS 60

QY 61 AAHCYKSRIOVRLGEYNIDVLEGNQFINSKAVIRHPNYSWILNDIMLIKSSPAVLN 120
 Db 61 AAHCYKRIQVRLGEHINVLGNQFVDSKILFHPNYSWTLNDIMLIKSSPVKLN 120

QY 121 ARVATISLPRACAAAGTQCLISGNGNTLSSTGNTYPELLQCLIDAPILTAQACEASYPGQIT 180
 Db 121 ARVATVALPSSCAPAGTQCLISGNGNTLSNGVNNPDLQCVDAPYLPQADCEASYPGKIT 180

QY 181 ENMICAGFLEGGKDCSGDGGPVVNCNGELQIVSWGCAQKPKGVYTKVCNFVDWIO 240
 Db 181 DNMCVGFLEGGKDCSGDGGPVVNCNGELQIVSWGCAQPDAPGVYTKVCNFVDWIO 240

QY 241 STIAAN 246

Db 241 DTIAAN 246

RESULT 5

TRRL1

trypsin (EC 3.4.21.4) I precursor - rat
N:Alternate names: trypsinogen I
C:Species: Rattus norvegicus (Norway rat)
C>Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 24-Sep-1999
C:Accession: B22657; A00948
R:Craik, C.S.; Choo, O.L.; Swift, G.H.; Quinto, C.; MacDonald, R.J.; Rutter, W.J.
J. Biol. Chem. 259, 14235-14264, 1984
A:Title: Structure of two related rat pancreatic trypsin genes.
A:Reference number: A22657; MUID:85054880; PMID:6094547
A:Accession: B22657
A:Molecule type: DNA
A:Residues: 1-246 <CRA>
A:Cross-references: GB:J00778; NID:g206507; PIDN:AAA98518.1; PID:g206508
A:Note: the authors translated the codon ATC for residue 6 as Leu and GAC for residue 17 as Ser.
J. Biol. Chem. 257, 9724-9732, 1982
A:Title: Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide sequences of two related rat pancreatic trypsin genes.
A:Reference number: A00948; MUID:82265624; PMID:6896710
A:Accession: A00948
A:Molecule type: mRNA
A:Residues: 1-246 <MAC>
A:Cross-references: GB:J00778; NID:g206507; PIDN:AAA98518.1; PID:g206508
C:Genetics:
A:Introns: 14/1; 67/2; 152/1; 197/3
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-23/Domain: activation peptide #status predicted <APT>
F:24-246/Product: trypsin I #status predicted <ENZ>
F:24-239/Domain: trypsin homology <TRY>
F:30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted
F:63,107,200/Active site: His, Asp, Ser #status predicted
F:75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 83.0%; Score 1101; DB 1; Length 246;
Best Local Similarity 80.5%; Pred. No. 2.2e-86;
Matches 198; Conservative 26; Mismatches 22; Indels 0; Gaps 0;

QY 1 MNPLLIILAFGLGAATVTPDDDDKIVGGYTCENSVYQVSLNAGYHFCGSLSDQWVVS 60
Db 1 MSALLIILVGAATVAFVLEDDDKIVGGYTCPEHSPVYQVSLNAGYHFCGSLINDQWVVS 60
61 AAHCYKSRIQVRLGEYNDVLEGNQFINSKVIHRHYNYSWILDNDIMLIKLSPPAVLN 120
Db 61 AAHCYKSRIQVRLGEHNVLEGEQFINSKVIHRHYNYSWLTNNNDIMLIKLSPPVKLN 120
QY 121 ARVATISLPACAPAGTQCLISGNGNTLSGGTNYPELLQCLDAPILTQAOCEASYPGQIT 180
Db 121 ARVAPVALPSACAPAGTQCLISGNGNTLSGNNPDLQCLVDPVLSQADCEAAYPGEIT 180
QY 181 ENMICAGFLGGKDCSCGDSGGPVVNCVNGELQIVSGYGCACQKNGPVYTKVCFNVDWIQ 240
Db 181 SSMICVFLGGKDCSCGDSGGPVVNCVNGELQIVSGYGCALPDNPVYTKVCFNVDWIQ 240
QY 241 STIAAN 246
Db 241 DTIAAN 246

RESULT 6

A27547

trypsin (EC 3.4.21.4) precursor, cationic|- rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 22-Jun-1999
C:Accession: A27547
R:Fletcher, T.S.; Alhadeff, M.; Craik, C.S.; Largman, C.
Biochemistry 26, 3081-3086, 1987

A:Title: Isolation and characterization of a cDNA encoding rat cationic trypsinogen.
A:Reference number: A27547; MUID:87271609; PMID:3607011

A:Accession: A27547

A:Molecule type: mRNA

A:Residues: 1-247 <FLE>

A:Cross-references: GB:M16624; NID:g206498; PIDN:AAA41985.1; PID:g206499

A:Superfamily: trypsin; trypsin homology

C:Keywords: calcium binding; hydrolase; protein digestion; serine proteinase

F:25-240/Domain: trypsin homology <TRY>

F:31-161,49-65,133-234,140-207,172-186/Disulfide bonds: #status predicted

F:64,108,201/Active site: His, Asp, Ser #status predicted

F:76,78,81,86/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 82.8%; Score 1098.5; DB 2; Length 247;
Best Local Similarity 78.5%; Pred. No. 3.6e-86;
Matches 194; Conservative 31; Mismatches 21; Indels 1; Gaps 1;

QY 1 MNPLLIILAFGLGAATVTPDDDDKIVGGYTCENSVYQVSLNAGYHFCGSLSDQWVVS 59
Db 1 MKALIFLAFGLGAVALPLDDDDKIVGGYTCQKNSLPYQVSLNAGYHFCGSLINSQWVVS 60
QY 60 SAHCYKSRIQVRLGEYNDVLEGNQFINSKVIHRHYNYSWILDNDIMLIKLSPPAVL 119
Db 61 SAHCYKSRIQVRLGEHNVLEGEQFIDAAKIIHRPSYNANTFNDIMLIKLSPPATL 120
QY 120 NARVATISLPACAPAGTQCLISGNGNTLSGGTNYPELLQCLDAPILTQAOCEASYPGQI 179
Db 121 NSRVTSVLSRCSGGTQCLVSGMGTLSGGTNYPSLLQCLDAPVLSCKSSYPGKI 180
QY 180 TENMICAGFLGGKDCSCGDSGGPVVNCVNGELQIVSGYGCACQKNGPVYTKVCFNVDWI 239
Db 181 TSNMFCFLGFLGGKDCSCGDSGGPVVNCVNGELQIVSGYGCACQKNGPVYTKVCFNVDWI 240
QY 240 STIAAN 246
Db 241 QQTVAAN 247

RESULT 7

B25852

trypsin (EC 3.4.21.4) II precursor [invalidated] - human
N:Alternate names: trypsin 2; trypsin, anionic; trypsinogen II
C:Species: Homo sapiens (man)
C>Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 08-Dec-2000
C:Accession: B25852; A61066; B43988
R:Emi, M.; Nakamura, Y.; Ogawa, M.; Yamamoto, T.; Nishide, T.; Matsubara, I.
Gene 41, 305-310, 1986
A:Title: Cloning, characterization and nucleotide sequences of two cDNAs encoding human trypsinogen II.
A:Reference number: A91544; MUID:86221712; PMID:3011602
A:Accession: B25852
A:Molecule type: mRNA
A:Residues: 1-247 <EMI>
A:Cross-references: GB:M27602; NID:g521217; PIDN:AAA61232.1; PID:g521218
R:Kimland, M.; Russick, C.; Marks, W.H.; Borgstrom, A.
Clin. Chim. Acta 184, 31-46, 1989
A:Title: Immunoreactive anionic and cationic trypsin in human serum.
A:Reference number: A61066; MUID:90091010; PMID:2598466
A:Accession: A61066
A:Molecule type: protein
A:Residues: 16-39, 'X', 41-42, 'XXXX', 47-49 <KIM>
R:Koivunen, E.; Huhtala, M.L.; Stenman, U.H.
J. Biol. Chem. 264, 14095-14099, 1989
A:Title: Human ovarian tumor-associated trypsin. Its purification and characterization.
A:Reference number: A43988; MUID:89340515; PMID:2503510
A:Accession: B43988
A:Molecule type: protein
A:Residues: 16-49 <KOI>
A:Experimental source: mucinous ovarian tumor cyst fluid
C:Genetics:
A:Gene: GDB:PRSS2; TRY2
A:Cross-references: GDB:335289; OMIM:601564
A:Map position: 7q35-7q35
A:Note: the human genome contains at least ten trypsin genes or pseudogenes, at least

C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-246/Product: trypsinogen II #status experimental <ZYM>
F:24-246/Product: trypsinogen II #status experimental <ZYM>
F:16-23/Domain: activation peptide #status predicted <APT>
F:24-246/Product: trypsin II #status predicted <TRY>
F:24-239/Domain: trypsin homology <TRY>
F:30-160,48-64,171-185,196-220/Disulfide bonds: #status predicted
F:63,107,200/Active site: His, Asp, Ser #status predicted
F:75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 82.7% Score 1097; DB 1; Length 247;
Best Local Similarity 81.8%; Pred. No. 4.8e-86;
Matches 202; Conservative 21; Mismatches 24; Indels 0; Gaps 0;

QY 1 MNPLLIILAFGAATVATDDDDKIVGGYTCENSPYQVSLNAGYHFCGSLISDQWVYS 60
DB 1 MNLLILITFVAALAAPFDDDDKIVGGYTCENSPYQVSLNAGYHFCGSLISDQWVYS 60
61 AACHYKSRIOVRLGEYNDIVLEGNQFINSKAVIRHPNYSWILNDIMLIKLSPPAVLN 120
61 AGCHYKSRIOVRLGEHNIIEVLEGNQFINAKIIRHPKYNSTLNDIMLIKLSPPAVLN 120
121 ARVATISLPACAPGTQCLISGNGNTLSSGTNTYPELLQCLDAPILTAQACEASYPGQIT 180
DB 121 SRVSAISLPATPAAGTESLISGNGNTLSSGADYPDELQCLDAPVLSQACEASYPGKIT 180
QY 181 ENMICAGLEGGKDCSCGDSGGPVVNCNGELQIVSWGYGCAQKNKPGVYTKVCFNFDWIQ 240
DB 181 NMFVGVLEGGKDCSCGDSGGPVVNSGELQIVSWGYGCAQKNRPGVYTKVYNFDWIK 240
QY 241 STTAANS 247
DB 241 DTIAANS 247

RESULT 8
A25852
trypsin (EC 3.4.21.4) I precursor [validated] - human
N:Alternate names: trypsin, cationic; trypsinogen I
C:Species: Homo sapiens (man)
C:Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 08-Dec-2000
C:Accession: A25852; B61066; A43988
R:Emi, M.; Nakamura, Y.; Ogawa, M.; Yamamoto, T.; Nishide, T.; Mori, T.; Matsubara, K.
Gene 41, 305-310, 1986
A:Title: Cloning, characterization and nucleotide sequences of two cDNAs encoding human
A:Reference number: A91544; MUID:86221712; PMID:3011602
A:Accession: A25852
Molecule type: mRNA
Residues: 1-247 <EMI>
A:Cross-references: GB:M22612; NID:g521215; PIDN:AAA61231.1; PID:g521216
R:Kjmland, M.; Russick, C.; Marks, W.H.; Borgstrom, A.
Clin. Chim. Acta 184, 31-46, 1989
A:Title: Human trypsinogen and cationic trypsin in human serum.
A:Reference number: A61066; MUID:90091010; PMID:2598466
A:Accession: B61066
Molecule type: protein
Residues: 16-43 <KIM>
R:Kjmland, M.; Russick, C.; Marks, W.H.; Borgstrom, A.
J. Biol. Chem. 264, 14095-14099, 1989
A:Title: Human ovarian tumor-associated trypsin. Its purification and characterization
A:Reference number: A43988; MUID:89340515; PMID:2503510
A:Accession: A43988
Molecule type: protein
Residues: 16-54 <KOI>
A:Experimental source: mucinous ovarian tumor cyst fluid
C:Genetics:
A:Gene: GDB:PRSS1; TRY1
A:Cross-references: GDB:119620; OMIM:276000
A:Map position: 7q35-7q35
A:Note: The human genome contains at least ten trypsin genes or pseudogenes, at least two
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; phosphoprotein; protein digestion; serine proteinase; z

F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-246/Product: trypsinogen I #status experimental <ZYM>
F:24-246/Product: trypsin I #status predicted <ENZ>
F:24-239/Domain: trypsin homology <TRY>
F:30-160,48-64,139-206,171-185,196-220/Disulfide bonds: #status predicted
F:63,107,200/Active site: His, Asp, Ser #status predicted
F:75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 81.8% Score 1085; DB 1; Length 247;
Best Local Similarity 80.2%; Pred. No. 5.1e-85;
Matches 198; Conservative 25; Mismatches 24; Indels 0; Gaps 0;

QY 1 MNPLLIILAFGAATVATDDDDKIVGGYTCENSPYQVSLNAGYHFCGSLISDQWVYS 60
DB 1 MNPLLIILFVAALAAPFDDDDKIVGGYTCENSPYQVSLNAGYHFCGSLISDQWVYS 60
QY 61 AACHYKSRIOVRLGEYNDIVLEGNQFINSKAVIRHPNYSWILNDIMLIKLSPPAVLN 120
DB 61 AGCHYKSRIOVRLGEHNIIEVLEGNQFINAKIIRHPQYDRKTLNNDIMLIKLSRAVIN 120
QY 121 ARVATISLPACAPGTQCLISGNGNTLSSGTNTYPELLQCLDAPILTAQACEASYPGQIT 180
DB 121 ARVSTISLPATPAAGTESLISGNGNTASSGADYPDELQCLDAPVLSQACEASYPGKIT 180
QY 181 ENMICAGLEGGKDCSCGDSGGPVVNCNGELQIVSWGYGCAQKNKPGVYTKVCFNFDWIQ 240
DB 181 NMFVGVLEGGKDCSCGDSGGPVVNCNGELQIVSWGYGCAQKNRPGVYTKVYNFDWIK 240
QY 241 STTAANS 247
DB 241 NTIAANS 247

RESULT 9
TRPGTR
trypsin (EC 3.4.21.4) precursor - pig (tentative sequence)
N:Contains: trypsinogen
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 24-Apr-1984 #sequence_revision 24-Apr-1984 #text_change 31-Mar-2000
C:Accession: A90641; A90368; A00947
R:Charles, M.; Rovey, M.; Guidoni, A.; Desnuelle, P.
Biochim. Biophys. Acta 69, 115-129, 1963
A:Title: Su le trypsinogene et la trypsin de porc.
A:Reference number: A90641
A:Accession: A90641
Molecule type: protein
Residues: 1-10 <CHA>
R:Hermanson, M.A.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.
Biochemistry 12, 3146-3153, 1973
A:Title: Determination of the amino acid sequence of porcine trypsin by sequenator ar
A:Reference number: A90368; MUID:73258692; PMID:4738933
A:Accession: A90368
Molecule type: protein
Residues: 9-231 <HER>
A:Note: at position 20, Ile and Val occur alternatively
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; polymorphism; protein digestion; serine proteinase;
F:1-23/Product: trypsinogen #status experimental <ZYM>
F:1-8/Domain: activation peptide #status experimental <APT>
F:9-23/Product: trypsin #status experimental <MAT>
F:9-224/Domain: trypsin homology <TRY>
F:15-145,33-49,117-218,124-191,156-170,181-205/Disulfide bonds: #status predicted
F:48,92,185/Active site: His, Asp, Ser #status predicted
F:60,62,65,70/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 80.7% Score 1071; DB 1; Length 231;
Best Local Similarity 83.0%; Pred. No. 7.3e-84;
Matches 191; Conservative 21; Mismatches 18; Indels 0; Gaps 0;

QY 17 PTDDDDKIVGGYTCENSPYQVSLNAGYHFCGSLISDQWVYSAACHYKSRIOVRLGEY 76
DB 2 PTDDDDKIVGGYTCENSPYQVSLNAGYHFCGSLISDQWVYSAACHYKSRIOVRLGEH 61

F:6-7/Cleavage site: Lys-Ile (enteropeptidase) #status experimental
F:13-143,31-47,115-216,122-189,154-168,179-203/Disulfide bonds: #status experimental
F:46,90,183/Active site: His, Asp, Ser #status experimental
F:58,60,63,68/Binding site: calcium (Glu, Asn, Val, Glu) #status experimental
F:131-132/cleavage site: Lys-Ser (autolytic) #status experimental

Query Match	73.7%	Score 978;	DB 1;	Length 229;
Best Local Similarity	77.2%	Pred. No. 6e-76;		
Matches 176;	Conservative 29;	Mismatches 23;	Indels 0;	Gaps 0;

QY	19	DDDDKIYGGYTCENSPYQVSLNAGYHFCGSLISDQWVYSAAHCKYSRIQVRLGEYNI	78
Db	2	DDDDKIYGGYTCGANTVPYQVSLNSGYHFCGSLINSQWVYSAAHCKYSRIQVRLGEDNI	61
QY	79	DVLEGNQFINSKVIHSPNYSWILNDIMLIKLSPPAVLNARVATISLPRAAPGTQ	138
Db	62	NVVEGNEQFISAKSIVHPSTNSNTLNDIMLIKLSAASLNSRVASISLPTSCASAGTQ	121
QY	139	CLISGWNLTLSSTNYPELLQCLDAPILTQAQCEASYPGQITENNICAGFLEGGKDCQG	198
b	122	CLISGWNLTSSGTSYDPDLKCLKAPILSDSSCKSAYPGQITSNWFCAGYLEGGKDCQG	181
QY	199	DSGGPVVNCGELQGIWVGSGCAQKNKPGVYTKVCFNFDWIQSTIAAN	246
Db	182	DSGGPVVCSKLGQIVSWGSGCAQKNKPGVYTKVCFNFDWIQSTIAAN	229

Search completed: June 20, 2003, 15:22:34
Job time : 41 secs